

AD-A101 339

ROYAL AIRCRAFT ESTABLISHMENT FARNBOROUGH (ENGLAND)
DIGITAL SPECTRAL ANALYSIS: A GUIDE BASED ON EXPERIENCE WITH AIR--ETC(U)
FEB 81 S L BUCKINGHAM

F/6 9/2

UNCLASSIFIED

RAE-TR-81014

DRIC-BR-78387

NL

100
200
300

END
DATE
FILED
7-8-81
DTIC

TR 81014

ADA10133

DTIC FILE COPY

UNLIMITED

14 RAE - TR-81014



LEVEL II

B7

ROYAL AIRCRAFT ESTABLISHMENT

*

9 Technical Report, b1014

11 February, 1981

DTIC
ELECTED

JUL 14 1981

S D

E

6 DIGITAL SPECTRAL ANALYSIS:
A GUIDE BASED ON EXPERIENCE WITH
AIRCRAFT VIBRATIONS.

by

10 S. L. Buckingham

12 597

*

Procurement Executive, Ministry of Defence
Farnborough, Hants

UNLIMITED

JCO

310450

81 6 29 046

UDC 519.246.87 : 620.178.53 : 534.13

ROYAL AIRCRAFT ESTABLISHMENT

Technical Report 81014

Received for printing 3 February 1981

DIGITAL SPECTRAL ANALYSIS: A GUIDE BASED ON
EXPERIENCE WITH AIRCRAFT VIBRATIONS

by

S. L. Buckingham

SUMMARY

→ A guide to digital spectral analysis is presented. The emphasis is on a practical engineering understanding of the techniques, based on experience gained in their application to the analysis of flight measurements of aircraft vibration in buffet. Consequently, particular attention is directed to the practical difficulties encountered when the duration of the available data is severely limited, and to the use of the coherence function as a tool in the interpretation of complicated responses. However, the presentation of the fundamental principles, and especially the inherent limitations of the techniques are relevant to any sphere of application.

Departmental Reference: FS 140

Copyright
©
Controller HMSO London
1981

LIST OF CONTENTS

	<u>Page</u>
1 INTRODUCTION	3
2 LIMITATIONS IMPLICIT IN PROCESSING DIGITAL DATA	4
2.1 Data sampling	4
2.2 Quantization and noise considerations	7
3 THE POWER SPECTRUM, AND ITS COMPUTATION VIA THE FAST FOURIER TRANSFORM	9
3.1 Elimination of very low frequency components	10
3.2 Windowing	11
3.3 Computation of Fourier transform and power spectrum	13
3.4 Averaging	16
3.5 Spectrum compensation	19
4 CROSS-SPECTRAL ANALYSIS	20
4.1 The cross spectrum	21
4.2 The coherence function	23
4.3 Non-symmetric cross-spectral functions	26
4.4 Compensation of cross-spectral functions	28
5 A CONVENIENT FORMAT FOR DISPLAYING PHASE AND COHERENCE INFORMATION	29
6 CONCLUDING REMARKS	30
Appendix A The convolution theorem	31
Illustration - Figure A1	32
Appendix B Demonstration of the spectral consequences of sampling	33
Illustrations - Figures B1 and B2	35
Appendix C Mathematical definition of Fourier transform and power spectrum	37
List of symbols	38
References	39
Illustrations	Figures 1-19
Report documentation page	inside back cover

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
<hr/>	
By	
Distribution	
Availability Codes	
Dist	Avail. to: <input type="checkbox"/> Special <input type="checkbox"/>
A	

1 INTRODUCTION

In recent years there has been a substantial increase in the use of digital spectral analysis techniques for an ever-widening range of engineering applications. The trend is certain to continue. The expanding use of the techniques is primarily a result of the development of the so-called fast Fourier transform algorithm, or FFT, which permits a very significant reduction in the cost (in terms of computing requirements) compared to earlier techniques. The ever-increasing speed, capacity, and availability of digital computers, the widespread introduction of digital instrumentation systems, and the advent of compact, purpose-built, FFT-based digital spectrum analysers have also contributed substantially to the increasing use of the techniques. Finally the continuing advances in the theoretical, and computational aspects of the method have both promoted, and been promoted by, the growing importance of the subject.

This Report presents a guide to digital spectral analysis, based on first-hand experience of the practical application of the technique to aircraft vibration data. The presentation differs from the currently available textbooks (*eg* Refs 1-4) in two respects. Firstly, the emphasis here is on a practical engineering understanding of the techniques, and not on rigorous mathematical proofs, which are readily available elsewhere if required. A sound understanding of the implications and particularly of the limitations of the techniques is essential if the results of such analysis are to be seen in perspective or if the techniques are to be applied successfully either on a general-purpose computer, or on a purpose-built digital spectrum analyser. Secondly, particular attention is devoted to the experience gained, and the refinements developed, in applying the techniques to the analysis of vibration measurements on a VC10 transport aircraft at buffet onset (the associated flight test investigation is reported in Ref 5). Although some of the refinements which were developed have been mentioned briefly in some of the most recent text books, the opportunity is taken here to supplement their advice and to highlight features that were of practical significance in a concise way that is not possible in the 'standard' texts.

In a number of respects, the present application was a severe test of the spectral techniques. The excitation of the structural vibration (*ie* the pressure fluctuation in the extensive separated flows) was complex, and the response of the airframe was extremely complex. There were some shortcomings in the instrumentation (principally low signal levels, and signal conditioning parameters which were a compromise between the requirements of several independent flight test programmes). In addition, practical difficulties in flying the test conditions accurately and steadily severely limited the duration of the data records. Nevertheless, the techniques described herein were applied successfully, and allowed the response of the aircraft to be understood in a qualitative (and, within certain limits, quantitative) way. No other currently available techniques could offer a more attractive approach to this problem than digital spectral analysis.

The material to be presented can be divided into three areas:

Data acquisition: Describing the implications and limitations of obtaining or converting data to a form suitable for digital analysis.

Computation of power spectra: Describing the complete process, from preparation of data into a suitable format, to applying the Fourier transform, and to computing the required spectrum.

Cross spectral analysis: Introducing the cross spectrum and coherence function as means of investigating the relationship between signals, presenting various additional cross-spectral parameters, and describing a format for displaying cross-spectral results.

In each area, the basic concepts and procedures are explained, and the practical experience of applying them is described where appropriate.

2 LIMITATIONS IMPLICIT IN PROCESSING DIGITAL DATA

The many significant advantages which can be realised by use of digital recording or processing systems are achieved at the cost of two inherent disadvantages:

- (a) The output represents a series of *samples* of the signal, and not a continuous record.
- (b) The digital output can only take one of a finite number of discrete values, *i.e.* it is quantized.

Both these aspects are fundamentally different from the kind of limitations experienced with analogue processing (*e.g.* frequency response and signal/noise ratio). The implications (for both time and frequency domain analysis) of these two facets of the digital approach are discussed below.

2.1 Data sampling

A digital recording system represents the recorded quantity by a series of samples. Although each sample may be a good representation of the signal at the instant of sampling, no information is recorded about the behaviour of the signal during the inter-sample periods. Simple interpolation procedures, or even the human eye can readily 'reconstruct' values, but, under some circumstances such reconstruction could be grossly misleading. Fig 1 shows the process of sampling and reconstruction for a hypothetical signal. In this case it can be seen that the reconstruction is a good representation of the original signal. However, Fig 2 shows an example where the reconstruction is entirely misleading. Clearly, the difference arises because of the higher frequency content of the signal in the second example. In the first example, there was never more than a small fraction of a cycle between adjacent samples; in the second case, there was more than half a cycle, and thus one or more maxima or minima between samples, but the interpolation procedure fitted the simplest possible curve between the data points, and hence reduced the apparent frequency of the signal. This process of introducing ambiguity into the frequency content of the signal is known as *aliasing*.

Strictly these limitations also apply whenever analogue data is transcribed into digital form even, for example, when trace records are read. Thus, the limitations can only be avoided by analogue processing of analogue records. However, in practice it was

the advent of digital recording, or automatic analogue-to-digital conversion which brought the problems to the fore.

Aliasing may be summed up in two statements:

(i) The recorded samples are always valid representations of the signal at the instant of sampling.

(ii) The absence of information about the behaviour of the signal between samples allows ambiguity in the interpretation of the samples. Hence any given set of samples could have originated from an infinite variety of original signals (for example, the samples shown in Fig 1c could have been generated by the signal shown in Fig 1a, or that shown in Fig 2a, or from suitable higher frequency signals, or from any linear combination of such signals).

For a limited number of applications (for example, calculation of means or rms levels of random signals) aliasing may not affect the integrity of the data, but if time domain signal-reconstruction, or spectral analysis is the objective, ways must be found to deduce the likely behaviour of the signal between samples.

The ambiguity can only be resolved if it is known *a priori* that the input signal was restricted to a sufficiently narrow frequency band, such that of all the frequencies which could have given rise to the observed samples only one could in fact have been present in the input signal (see Fig 3). In practice, this is usually achieved by using a low-pass pre-sampling filter to restrict the bandwidth of the signal before it is sampled. The design of pre-sampling ('anti-aliasing') filters, and their relationship with optimum sample rates and interpolation processes for time domain signal-reconstruction of signals with flat, broadband spectra is covered by Gardenhire⁶. In this Report attention is focussed on frequency domain analysis of signals which may have very 'peaky' spectra.

The effect of aliasing on the spectral content of signals is as shown in Fig 4* **. Specifically, the spectrum of the sampled data consists of the sum of the spectra of many 'images' of the spectrum of the original (continuous) data. The images appear at equally spaced frequency intervals throughout the entire frequency axis ($-\infty$ to ∞). The spacing of the images is equal to the sample rate which was used. As can be seen in Fig 4, the true spectrum, and the first aliased image intersect at a frequency equal to half the sample rate. This frequency is generally referred to as the Nyquist frequency. Thus any significant signals at frequency above the Nyquist frequency will have aliased images which contribute power into the frequency band from dc to the Nyquist frequency.

* This result can be derived by means of the convolution theorem. This theorem is described in Appendix A, and used to demonstrate the spectral consequences of sampling in Appendix B.

** Fig 4 shows the spectra as two-sided, *i.e.* with components at both positive and negative frequencies. Although negative frequencies are meaningless in the real world, and actual measured spectra are always shown for positive frequencies only, the two-sided representation used here is widely used for mathematical convenience when considering manipulations of spectra. The equivalence between the two formats is discussed in the standard text books; for the present Report, it is sufficient to accept their equivalence. Note that the two-sided spectrum of a real signal will always be symmetric.

In principle there are three methods which can be used to combat aliasing. Firstly, it may often be desirable to filter the signal before digitization, so that no significant power remains at frequencies beyond the Nyquist frequency. The filtering may be achieved within the signal conditioning stages, or by selecting transducers which effectively act as low-pass filters by virtue of their limited frequency response. High-order filters are commonly used in such applications, to ensure a positive cut-off of unwanted signals, combined with minimum attenuation of signals within the frequency band of interest. It is important, however, that the benefits of filtering are carefully balanced against the disadvantages, both in terms of cost and complexity, and in terms of possible other penalties. For example, if powerful high-frequency signals are present, they may saturate the transducer, and hence corrupt its output at all frequencies. If such saturation were masked by the anti-aliasing filters, very severe errors would arise. Furthermore, high order filters will introduce large phase distortions even at low frequencies. Although this is not usually a disadvantage in spectral analysis (because the distortion can be corrected within the analysis if necessary), if the same data records are to be used for time-domain analysis also, high order filters may be unacceptable unless care is taken to use filters with linear phase/frequency characteristics. Indeed, the conflicting requirements of two experimental programmes led to low order filters being used both for accelerometer and strain gauge signals during the larger part of the VC 10 buffet research programme.⁵.

Secondly, the higher the sampling rate, the less likely it is that aliased signals will cause significant interference, provided the spectrum of the (continuous) data does not contain major peaks at very high frequencies. The beneficial effects of increasing the separation between the spectral images (*ie* of increasing the sampling rate) is demonstrated in Appendix B and Figs 4 and 5. There are, however, disadvantages in using excessive sample rates. The most obvious is that higher sampling rates need increased capacity within the recording system (and also, possibly, within the data handling and analysis system). It will be shown later that increasing the sample rate can never improve the resolution of the computed spectra. Indeed, if the capacity of the analysis computer limits the number of samples which can be used to compute a spectrum (the block size), then increasing the sample rate will degrade the spectral resolution. In some applications, this disadvantage is compensated by the fact that less time is required to acquire a given number of data values. These trade-offs are examined in greater detail in section 3.3.

A third technique can be applied to prevent high frequency narrow band signals aliasing into the frequency band of interest, provided that the frequency of any such signals is constant, and is known in advance. A typical example might be the 400 Hz ac power frequency used on aircraft. Since for a given sample rate, the frequency to which a high frequency signal will alias is fixed, this can be used as a constraint on the sample rate so that the aliased frequency is well separated from frequencies of interest. This can usually be achieved by a very modest change of sample rate; however, with many recording systems sample rates are limited to a number of fixed values (often in steps of binary multiples of the lowest system sample rate), and are thus not amenable to this approach.

It is important to note that all three of the above techniques can only be used effectively if there is some knowledge of what high frequency signals may be present, and if the 'bandwidth of interest' is well established before the recording system is set up*. The necessary information may be known from past experience, or theoretical considerations. Otherwise it will be necessary to conduct pilot experiments (possibly using very high sample rates, or even using analogue analysis), or to proceed on the basis of estimates, and to monitor results closely to ensure that serious aliasing is not occurring. This in turn can only be achieved if analysis software, and data handling utilities are available and tested before actual data recording starts. Monitoring may take the form of comparing results obtained using differing filters or slightly different sample rates. In the latter case, any aliased peaks will be revealed by a change in apparent frequency, which is equal to some integer (usually 1, because the first image spectrum is the most likely source of aliases) times the change in sample rate**.

The problems of aliasing are thus inherent in the idea of sampled data, and a thorough understanding of the phenomena, and a conscious effort to detect and avoid aliasing is essential if meaningful results are to be obtained from spectral analysis. In the future, the continuing trend towards increased capacity in recording systems and in the computers used for spectral analysis will doubtless mean that it will become increasingly attractive to raise sampling rate as a primary method of combating aliasing. Another alternative could be the development of recording and analysis systems based on non-periodic data sampling. Some initial investigations into such an approach have revealed some promise, but much further work will be required before it will be known if the complexity of such an approach can be justified in any but the most extreme circumstances.

2.2 Quantization and noise considerations

At any instant in time, analogue records can take on any one of an infinite number of values in a continuous, but bounded range. The actual value retrieved will not, in general, correspond exactly to the value of the parameter which was recorded because some random noise will have been added to the true signal during transduction, transmission, recording, storage, and replay (and also because of frequency response characteristics of the whole system). Each possible noise source will have its own distinctive noise spectrum, and in addition the noise signals injected in the earlier stages of the sequence of operations from transduction to replay will be modified according to the frequency response characteristics of the later stages.

When a signal is digitized, an additional, unique, type of noise is inflicted on it. At any instant in time, a digital record can only take one of a finite number of possible discrete values in a bounded range. Thus, if a digital record is hypothetically compared

* Indeed one of the major difficulties in the VC 10 buffet research programme was that the range of frequencies found to be important was substantially higher than had been expected. As a result, margins from aliasing were considerably smaller than the ideal.

** This method cannot be applied to data systems where all possible sample rates are binary multiples of a basic rate, because doubling the sample rate does not in general change the apparent frequency of an aliased narrow band signal.

with the exact value of the parameter which was recorded, an additional source of error is revealed; the difference between the 'true' value, and the closest discrete digital value. This phenomenon is illustrated in Fig 6.

Provided that the signal amplitude is large compared to the spacing of the discrete digital signal levels, these so-called quantization errors will be random, and will not be correlated from sample to sample. The quantization error for each sample will always be within the range $-\frac{1}{2}$ to $+\frac{1}{2}$ of the quantization interval, and any value within the range has an equal probability of occurring. As a result the effect of quantization can be shown to be equivalent to a pure white noise signal, with rms level equal to $1/\sqrt{12}$ times the quantization interval. As the quantization noise is inflicted during digitization, it is completely unaltered by any analogue filters which may be used in signal conditioning. However, it is worth noting that at present virtually all practical digital acquisition systems are technically hybrid systems, since the transducers usually produce an analogue output which is transmitted to the recording system before being digitized. The analogue stages of such a system will, of course, be subject to analogue noise consideration in the usual way. The advantage of digital recording is that once the signal has been recorded, no further noise should normally be added during storage or replay, although in practice, 'drop-out' on the data tape may sometimes occur. The majority of drop-outs will be detected by the parity-checking features of the recording system. Although some recent recording systems incorporate limited data redundancy, so that corrupt samples can be reconstructed, in most cases the analysis software will be required to ignore any corrupt samples or overwrite them with 'reasonable' values. Occasionally there may be significant numbers of corrupt samples which are not detected by the parity check system. Indeed, it is believed that on occasions transients within the signal conditioning and analogue-to-digital conversion process gave rise to significant data corruption during the VC 10 buffet research programme.

One method for detecting such data is to compare each data value with the mean data value for the block, expressing the difference in terms of the standard deviation of the data values within the block. If the value is greater than some threshold, the sample should be rejected. For a steady pure-tone sinusoid, the maximum difference would be $\sqrt{2}$ (the ratio of peak value to rms for a sinusoid). For random vibration data, rather larger values will be found. If the threshold is set too low, then some genuine data may be rejected; or if it is too high, corrupt values will be accepted. For the VC 10 buffet analysis, a value of 5 was chosen, with a block size of 256 samples. The probability of a block of 256 samples containing a genuine value exceeding 5 standard deviations is negligible, while accepting one 5-sigma corrupt value will only increase signal rms levels by 5%, which will not obscure significant spectral peaks. If corrupt values are detected, they should be eliminated, the standard deviation of the data recalculated, and the data checked through again with the corrected value. If corrupt values are not corrected, or are inadequately corrected, they will manifest themselves as spikes in the recorded signal. A single such spike has a flat, broadband spectrum, which, like the quantization noise, is unaffected by analogue signal conditioning because it occurs after digitization. Multiple spikes can produce complex harmonic components in computed spectra.

Although many alternative interpolation procedures are possible to compute values to substitute for rejected data values, the very simple approach of repeating the preceding value was found to be adequate for the present work.

For the majority of modern applications, the advantages of digital recording and analysis greatly outweigh the difficulties outlined above, provided that the appropriate measures are taken to control aliasing, and that the limitations implicit in the techniques are understood.

3 THE POWER SPECTRUM, AND ITS COMPUTATION VIA THE FAST FOURIER TRANSFORM

The power spectrum, or power spectral density (PSD) function is basically a very simple concept that is widely used and understood. Accordingly, only a brief resumé of its most significant features is appropriate here; a brief mathematical definition of the power spectrum, and its relationship to the Fourier transform is given in Appendix C. More detailed presentations can be found in Refs 1-4 and 7. The power spectrum describes the way in which the power in a signal is distributed across a range of frequencies. It contains no information about the phases of the various components of the signal; indeed when one continuing signal is seen in isolation there is in general no genuine significant reference against which phase can be defined or measured. Since the function defines the power density as a function of frequency, the total power within a particular frequency band can be obtained by integrating the PSD with respect to frequency between the required frequency limits.

The currently preferred method of digital spectral analysis is to use the fast Fourier transform to calculate the transform of the data, and the power spectrum can then be calculated as the square of the modulus of the transform. Note that the Fourier transform, and hence the PSD, of a non-periodic signal is a continuous function of frequency, irrespective of whether the original signal is continuous or sampled. However, the FFT treats a duration-limited record of a non-periodic signal as though it is one whole cycle of a complex periodic waveform. Hence it evaluates a 'discrete' transform (and thus defines the 'discrete' power-spectrum) at the specific frequencies which are harmonics of the record duration. These values are then effectively seen as samples, in the frequency domain, of the continuous spectrum which would correspond to the original, non-periodic data.

There are in general up to five stages in the calculation of reliable power spectra from digital data records:

- (a) Elimination of very low frequency components.
- (b) Windowing.
- (c) Computation of Fourier transform, and hence power spectrum.
- (d) Averaging to enhance statistical accuracy.
- (e) Compensation of computed spectra for known system response characteristics.

These stages are discussed in detail in the following sub-sections.

3.1 Elimination of very low frequency components

Very low frequency signals may occur because of undesirable transducer drift (particularly common in strain gauges), or as an inevitable consequence of the manner in which an experiment has to be carried out (for example, during flight vibration testing, vertical accelerometers will pick up not only vibration, but also rigid body accelerations caused by variations of the aircraft flight path).

When the period of such components of whatever origin is greater than the block* duration, it is important that steps are taken to eliminate them. Serious distortion of the low-frequency end of the computed spectrum may otherwise occur, because the analysis procedure implicitly represents the signal as being composed entirely of components with periods not longer than the block duration.

In some applications it may be possible to avoid or minimize such drifts by analogue filtering, or by suitable selection of transducers, but in many applications, such 'drift' must be eliminated during pre-processing. Although in principle if sufficient length of data is available these low frequency components could be removed by means of a high-pass digital filter, with a low cut-off frequency, in practice the very low frequencies involved and the high sample rates make such an approach unattractive. Instead it is usual to look at the blocks of the data individually, and to remove drift by subtracting a 'drift line' from the data. The 'drift line' can be calculated by a least squares method (Fig 7). However, this method must be used with considerable care because in some circumstances, genuine moderate frequency signals in the data can give rise to apparent drifts. This is illustrated in Fig 8. Here data with a genuine moderate frequency signal (but no drift) will give a calculated drift line indicating a spurious drift. When this line is subtracted from the data, the spectrum will show spurious power at other low frequencies, and a loss in power at the true signal frequency. Calibration signals can be particularly prone to this effect. Thus drift elimination should only be used when drift is known to be present.

One possible method to distinguish between the genuine very low frequency signals (which will need to be eliminated), and phasing effects in moderate frequency signals (which should not) could be based on the discontinuity of the drift lines from block to block. The procedure is illustrated in Fig 9. Where the discontinuity from block to block was small in comparison with the rms level of the signal after the drift is removed, it would be taken as a genuine trend.

Such a technique would have the disadvantage of requiring the evaluation of drift lines for several blocks of data before analysis could start. This could be an undesirable complication for analysis procedures based on data from magnetic tape, and would not be acceptable for real-time spectral analysis procedures. It was not employed directly during analysis of VC 10 buffet data, although mean values and drifts for each block were tabulated, so that checks could be made manually if circumstances demanded.

* A 'block' in this context contains the number of samples which will be used to compute a spectrum whether via the FFT or otherwise. Selection of block size is discussed in section 3.3.

3.2 Windowing

Any practical dynamics experiment can only produce a data stream of finite duration. Statistical considerations generally dictate that the data is then sub-divided into a series of 'blocks' prior to analysis. Effectively analysis techniques then treat each block of data as though it represents one cycle of a complex waveform, which continues to be repeated for all time, as shown in Fig 10a. If in fact the block contains only a signal, or combination of signals, which are exactly periodic within the block duration, no errors will occur (Fig 10b). However, any non-periodic components will be subject to abrupt transients at the end of each block (Fig 10c), and thus will become distorted by the analysis. Because the last samples of the block are treated as though they are followed by (a repeat of) the first samples of the block, the process of analysis is sometimes described as cyclic. A more precise understanding of the way non-periodic frequencies are distorted can be developed by the application of the convolution theorem (see Appendix A).

In effect the process of analysing a block of data of finite duration is equivalent to analysing a signal which is the product of two signals; one of the two being a data stream of infinite duration, and the other being a weighting or 'window' function, which has the value 1 during the interval when data is actually available, and has the value 0 at all other times. This interpretation is illustrated in Fig 11. This particular rectangular window is known as the boxcar window. The boxcar window is simple to apply, and at first sight its constant weighting appears the obvious way to segment the data. However, if the convolution theorem mentioned above is used to investigate the properties of the boxcar window, it becomes apparent that alternative windowing procedures, involving non-constant weightings can be preferable. Fig 12 shows a comparison of several window functions, in both time and frequency domains. The ideal window would appear to be one whose spectrum is as shown in Fig 12a - with one infinitely narrow peak at zero frequency. If such a function were convolved with the data spectrum, the resulting spectrum would be identical to the data spectrum, and no 'blurring' would have occurred. However, when viewed in the time domain, it is immediately obvious that such a window cannot be used in practice, because the corresponding time domain representation is a continuous dc level. Obviously multiplying a signal by a steady dc level does not affect the spectral content, but equally obviously such an approach does not produce a time history which is restricted to a finite duration. Thus such a 'perfect' window cannot be reconciled with a finite data duration.

The box-car window is shown in Fig 12b. The spectrum of such a window of duration τ has the form $|\tau \sin \pi f\tau / \pi f\tau|^2$. Clearly such a window differs significantly from the ideal form shown in Fig 12a. When this window's spectrum is convolved with the 'true data spectrum, the result will show considerable blurring of any narrow spectral peaks, because of the width of the major lobe of the window spectrum. Furthermore, the side-lobes have an amplitude which is a significant proportion of the major lobe amplitude over a substantial frequency band. This means that any major peaks in the true data spectrum will contribute apparent power over a wide range of frequencies in the computed spectrum. This effect is called leakage. Leakage is discussed in detail in Ref 8.

Numerous attempts have been made to design windows which have sidelobes which decay more rapidly than those of the boxcar window. The simplest of these is the Hann, or cosine bell window, shown in Fig 12c. In this case, a weight of $\frac{1}{2}(1 + \cos 2\pi(t/\tau))$ is applied to the data values over the range $-\tau/2$ to $\tau/2$. The transform of the window has a major lobe which is twice as wide as that of the boxcar window, and thus is worse from the point of view of blurring; but the sidelobes decay much more rapidly (in fact they decay as f^{-3} , rather than as f^{-1} in the case of the boxcar window), so that mutual interference between well separated peaks is greatly reduced.

In an attempt to retain the narrow main lobe of the boxcar window, while achieving the favourable sidelobe decay of the Hann window, Ref 9 advocates the data window shown in Fig 12d. Here, the initial and final 10% of the boxcar window are replaced by half cosine bells. Although the sidelobes of this window do decay more rapidly than those of the boxcar, in practice this advantage becomes evident only at frequency offsets beyond about $5/\tau$. Notice that all these 'improved' window functions apply low weightings to the beginning and end of each block, and hence smooth out the inter-block discontinuities which were shown in Fig 10c.

Other window functions, and some criteria for assessing their performance, are given in Refs 10-15. It is clear that there is no generally accepted best window but rather that the different advantages and disadvantages of the various windows should be considered in the light of the requirements of the particular experiment. In the case of the VC 10 buffet programme, the 10% cosine taper window was used in the initial power spectrum calculations (at least in part in direct consequence of the successful use of this window in the earlier analysis of wind tunnel tests); but the Hann window was applied in the more recent cross-spectrum calculations.

Before leaving the subject of windowing, three further points should be made. Firstly, the application of any window other than the boxcar has the effect of attenuating the data. Thus, it is important that suitable factors are applied to calculated power spectra to compensate. The appropriate factor can be shown to be the inverse of the mean squared value of the window function used^{1,11}. Secondly, except in the case of the boxcar window, the windowing process applies unequal weights to samples which are actually of equal statistical validity. This has the effect of reducing the effective number of degrees of freedom of the data, and thus has an adverse effect on the statistical stability of the spectrum estimates. This point is considered briefly in Ref 11. Finally, any practical window has a finite duration, and thus may be imagined as the product of a (periodic) shaping function, and a duration-limiting boxcar function. Thus windowed data can be thought of as the product not of two functions (continuous data and a shaped, duration-limited window), but of three (continuous data, duration-limited boxcar window, and a periodic shaping function). Hence the transform of shaped, duration-limited data may be calculated alternatively by convolving the transform of the periodic shaping function with the transform of the duration-limited signal. If the transform of the shaping function is complicated, this approach is not attractive. However, the transform of the periodic cosine bell weighting function is very simple,

comprising only three coefficients of value $\frac{1}{3}$, $\frac{1}{3}$ and $\frac{1}{3}$ at frequencies $-1/\tau$, 0 and $+1/\tau$. As these three frequencies are at the same spacing as the spectral lines produced by the FFT (as described in the next section), the convolution process can be carried out simply by replacing each transformed value by a suitable weighted combination of three adjacent values:

$$X'(k) = \frac{1}{3}X(k-1) + \frac{1}{3}X(k) + \frac{1}{3}X(k+1) . *$$

The choice between using a shaped window directly, or of using a boxcar window and applying an effective shaping by processing the transformed data as shown above is thus purely a matter of computational convenience for the (admittedly small) family of window shapes which have a simple transform. It should be stressed that just as the shaped window is applied to the data amplitudes, the spectral smoothing described here must be applied to the (complex) Fourier transform and not directly to the power-spectrum. Refs 1, 9 and 10 discuss the approach in more detail. Refs 16 and 17 conclude that the implementation of windows by weighting in the frequency domain is often the preferable method.

3.3 Computation of Fourier transform and power spectrum

Although it is possible to compute power spectra by analogue processing, or digitally either by direct numerical application of the Fourier integral, or by means of auto-correlation techniques, the Fast Fourier Transform (or FFT) has now won almost universal acceptance as the preferred technique for most applications of spectral analysis. As its name suggests, the FFT is used because it offers substantial advantages in speed of computation compared to other digital techniques. The actual speed ratio increases with the number of samples to be used to calculate the spectrum. Direct Fourier analysis requires of the order of N^2 operations (where an operation comprises a complex multiplication and addition) for a block size of N samples. Using the auto-correlation technique, this can be reduced to Nm operations if m lag values are used. The FFT can require as few as $4N \log_2 N$ operations**. The FFT achieves this advantage by calculating the results of many short transforms, and combining these to form the overall transform. Although theoretically this short cut is available wherever the block size is not prime (eg a block size of 26 could be analysed as two blocks of size 13, and 13 blocks of size 2, and the results combined to generate the spectrum of the full block), in practice the greatest saving in computational time occurs when $N = 2^P$, where P is any integer, and this is the basis used in speed comparison above. As a result, the name 'FFT' is generally taken to refer to this special case. The detailed derivation of the means by which the many short spectra are calculated and combined is beyond the scope of this Report. For these details and the resulting algorithms the reader is referred to one of the standard textbooks, eg Ref 1. Examples of implementation of the algorithm in a high level language may be found in Refs 4, 18 and 19.

* As shown, the expression relates to a boxcar window from $-\tau/2$ to $+\tau/2$. If the window used is in fact $0-\tau$, the sign of the terms in $k+1$ and $k-1$ should be changed.

** Subject to the condition that $\log_2 N$ is integer.

The speed advantage of the FFT is achieved at the cost of some loss of flexibility, both in terms of block size (as already mentioned) and also in spectral resolution.

Using the following notation:

S sample rate (per second)
 N block size (*ie* number of data points used to calculate one spectrum)
 (must be a power of 2)
 ℓ number of frequencies at which spectrum is evaluated
 T_B duration of data to produce one spectrum
 f_{\min} frequency of lowest-frequency spectral estimate
 f_{\max} frequency of highest-frequency spectral estimate
 δf spacing of spectral estimates,

the following relationships are found for the basic FFT:

$$T_B = \frac{N}{S}$$

$$\ell = \frac{N}{2}$$

$$f_{\max} = \frac{S}{2}$$

$$f_{\min} = 0$$

$$\delta f = \frac{f_{\max} - f_{\min}}{\ell} = \frac{S}{N} = \frac{1}{T_B} ,$$

since the FFT always produces $N/2$ spectral estimates at equal frequency increments from 0 to $S/2$ Hz. Note in particular that the spacing of the spectral estimates, δf , is independent of the sample rate if the block time T_B is fixed, or is proportional to sample rate (*ie* higher sample rate produces estimates at wider spacing) if the block size, N , is fixed. Hence increasing the sample rate can never contribute to better spectral resolution. When circumstances demand, two techniques may be used to improve the spectral resolution. If r zeros are added to the data string before analysis, then the effective block length becomes $N' = r + N$ (and now it is N' , not N which is constrained to be a power of 2), the relationships become:

$$T_B = \frac{N}{S} \quad (\text{unaltered})$$

$$\ell = \frac{r + N}{2}$$

$$f_{\max} = \frac{S}{2} \quad (\text{unaltered})$$

$$\delta f = \frac{2}{r + N} .$$

Thus the same band of frequencies is covered, but more estimates are produced, giving better resolution. Note, however, that the larger block size increases the time required

for analysis. Furthermore, with this method, N data values are being used to calculate power and phase values at $(N+r)/2$ frequency points: thus the $(N+r)$ numbers which represent the phase and power information are not all independent. In other words, from the point of view of information content, this process is only equivalent to computing the basic $(N/2)$ point spectrum, and then obtaining values at other intermediate frequencies by means of a complicated (non-linear) interpolation procedure.

In some applications these weaknesses can be avoided by using instead the second method of enhancing frequency resolution, which is commonly referred to as the zoom FFT. In this method, the frequency band of interest is deliberately aliased into a low frequency band, and then processed in the normal manner. The aliasing can be achieved either during recording (by using analogue band-pass filters to reject all data outside the required frequency band, and then by using a low sample rate), or during analysis by recording 'normally', then using a digital band-pass filter to limit the bandwidth, and then reducing the effective sample rate by systematically omitting samples (decimation). While the first of these methods has the advantage of simplifying analysis, and reducing data recording and storage requirements, it has the disadvantage that only the aliased data is recorded, and there is no possibility to modify the analysis strategy in the light of experience (except by re-running the experiment). The sample rate is chosen so that when it is multiplied by some integer, the result is the lowest frequency of interest. The relationships between the various processing parameters then become as follows: (notation as on opposite page except where indicated):

Z integer factor relating f_{\min} to sample rate
 δf frequency increment of basic FFT on range 0 to f_{\max} ,

$$s = \frac{f_{\min}}{Z}$$

$$\delta f_{\text{zoom}} = \frac{s}{N} = \frac{f_{\min}}{NZ} = \frac{2f_{\max}}{(2Z+1)N} = \frac{\delta f}{2Z+1}$$

$$l = \frac{N}{2}$$

$$f_{\max} = \left(\frac{N}{2} \delta f_{\text{zoom}}\right) + f_{\min} = f_{\min} \left(1 + \frac{1}{2Z}\right)$$

$$T_B = \frac{N}{s} = \frac{1}{\delta f_{\text{zoom}}} .$$

Thus the zoom transform allows all of the available spectral lines to be used to describe the spectrum in the narrow band $f_{\min} - f_{\max}$. Its advantages are thus clearly most pronounced when $f_{\min} \gg f_{\max} - f_{\min}$. Note that like the basic FFT, the frequency resolution is equal to the inverse of the block time, but the reduced sample rate means a given block size corresponds to a longer block time.

3.4 Averaging

It was shown above that the FFT uses N samples of data to compute phase and amplitude information of $N/2$ frequency points. Thus the number of independent parameters calculated by the procedure is exactly equal to the number of input variables. Furthermore each of the inputs is equally valid, and each of the outputs depends on each of the inputs to the same degree. As a result it will be appreciated that the statistical variability of the calculated parameters will be as wide as the variability of the input variables. Thus when the FFT is used to analyse random data, it will be necessary to take average values over a number (m) of computed spectra in order to obtain statistically stable results. It is shown in Ref 1 that (for the boxcar window) if results are averaged over m independent spectra, the normalised standard error (*i.e* the standard deviation of the population of estimates of the power, normalized by the true power) at each frequency is

$$\epsilon = \frac{1}{\sqrt{m}} .$$

Thus if no averaging is done ($m=1$), the calculated power at any given frequency will not be the true power, but rather it will be one value from a population of values, and the population will have a mean, and standard deviation equal to the true value.

Alternatively, it would be necessary to average 100 spectra if the averaged power is required to be a member of a population whose mean is the true value, and whose standard deviation is 10% of the true value.

It may at first appear that the above argument necessitates averaging in the case of random data but that no averaging would be required with deterministic data. However it is shown in Ref 10, that even in the case of deterministic data, averaging is required. The requirement arises because in general there will be a number of complete cycles, plus a part of a cycle within the duration of the data used to compute a spectrum. The energy within the part-cycle will depend on the phase of the signal. Thus spectra computed from different blocks of data will have different power levels (because the data and the block are not synchronized, so the phase difference is random), and the true power level is the average of all possible phase relationships. However, this effect is only powerful at low frequencies; for signals with more than about ten cycles within the block duration the fluctuation will be insignificant in practical terms.

In the real world, of course, spectral analysis is frequently applied to problems which cannot be described as truly random, or as totally deterministic. In such cases the degree of averaging required can only be determined on the basis of experience. For example, when a lightly-damped structure is subjected to random excitation, the energy stored in the vibrations will be far greater than the work done by the excitation in one cycle. The intensity of the response will tend to vary less than the intensity of the random excitation varies, so less averaging may be required. The objective of the VC 10 buffet programme, was the identification of the active modes of vibration. For this it was found that data (at constant flight condition) within a 20 second run were quite consistent, and so the average of ten, 2-second blocks was the standard analysis

procedure*. On some occasions, the pilot was unable to maintain steady conditions for 20 seconds, and it was thus necessary to average over fewer blocks. Although the results obtained from shorter records were treated with less confidence than results from longer runs, the general results were usually in acceptable agreement. On a few occasions steady conditions were maintained for substantially longer than 20 seconds. These runs provided a useful standard against which to view the stability of the 20 second runs, and served to confirm that the average of just ten spectra was an acceptable estimate. On the other hand, comparison of data from different flights, at nominally the same flight conditions did reveal significant differences in response. Although such differences could have been seen as the result of too low a number of averages, it was felt that slight differences in test conditions (eg fuel weight and distribution) were a more probable cause.

The requirement for averaging thus can cause a very considerable increase in the duration of data required, possibly by more than two orders of magnitude for true random data. In many experiments, there are practical limitations on the possible duration of recording. In the case of the VC 10 buffet programme, the aircrew found extreme difficulty in maintaining steady condition for longer than 20 seconds, particularly at certain Mach numbers. In other examples the duration of the actual phenomena may be limited (eg spacecraft launch vibration). The question inevitably arises as to whether there are any means of reducing the volume of data required. In fact, a few techniques are available but which, if any, are suitable for a particular application depends on the nature and requirements of the experiment.

Firstly we can consider the parameters relating to the volume of data required to compute each spectrum. The relationships between these parameters were stated in the previous section. If zero padding is not used, then the necessary data duration for each spectrum is the inverse of the frequency spacing of the spectral estimates (whether or not the zoom transform is used) and is independent of sample rate and number of samples per spectrum. Thus the duration of each 'block' can only be reduced either at the expense of a coarser frequency scale, or by resorting to zero padding. In practice, such tactics may permit some reduction in duration, though major reductions in duration are unlikely.

An entirely different approach is to obtain more data, and hence more spectra, by repeating the experiment several times. In some cases this may offer a very cost effective way of obtaining greater statistical reliability of spectra, but in other cases, the cost of repeat experiments may be totally prohibitive. Furthermore in some circumstances it may not be possible to do comparable repeat experiments. For example, if vibration data for two nominally identical structures were compared, it might be found that the frequencies of one of the dominant vibrations differed slightly, perhaps because of minor manufacturing differences. In this case, the repeat experiment might well be valuable because it demonstrated the variability; nevertheless it would be wrong

* This is in reasonable agreement with the averaging procedures developed for a study of vehicle vibrations²⁴.

to attempt to average the two sets of spectra directly, because this would lead to the vibration in question being represented as having two distinct frequencies (Fig 13)*.

Finally, the technique known as overlap processing can be applied in order to ensure that the information contained within a data record is extracted as effectively as possible. Depending on the degree of overlap employed, there are two possible methods by which overlap processing can improve the statistical stability of a spectrum. There is also a fundamental limit to the degree to which statistical stability can be enhanced.

When a long data record is analysed as a series of blocks, using a boxcar window, and no overlap, all the data is taken into consideration with an equal weighting (Fig 14a). However, if the boxcar window is replaced by any other window (either directly, or indirectly as discussed in section 3.2 above), the data do not receive equal weightings, and indeed some samples at the very end of each block have such a low weighting as to be virtually ignored (Fig 14b). Depending on the exact window shape, the various data points can be accorded weightings which are equal, or almost equal, if the appropriate degree of overlap is applied. Thus Fig 14c shows a 90% cosine taper window applied with 10% overlap, and Fig 15d shows the Hann window applied with 50% overlap. In both these cases the result is that a weighting of unity is effectively applied to all but the very first and last data points. Such an approach to overlap processing is investigated in Ref 20. Not all data windows can be made to overlap exactly, as these cosine-based windows can, but where necessary, an appropriate degree of overlap can be estimated, so that the combined weighting remains close to unity.

An alternative approach might be to employ a very high percentage of overlap (eg 90% as in Fig 14e). In a sense this approach makes the maximum use of the information in the data, albeit at the expense of a very substantial decrease in processing speed. For example, in Fig 15a (no overlap) the sample at X is analysed in the context of the signal between points A and B only. With 50% overlap (Fig 15b), X is analysed in the context of the signal between A and B, and between C and D. Thus it is in fact compared to data between times A and D. In the limiting case of very high overlap, the signal at X is seen in the context of signal between P and Q, where P occurs one block duration before X, and Q one duration after X. It is apparent that in the limit, this procedure is analogous to auto-correlation analysis. Clearly overlap processing allows the phase relationship between deterministic signals, and block boundaries to be varied, and thus is an effective way to obtain averages for such data. However, it was noted above that such a procedure was only necessary in the case of low frequency signals.

The use of overlap processing allows the available data to be scrutinized more closely than is possible without overlapping. Although this is beneficial, in one important respect it is not really a substitute for longer data duration when analysing random data. By definition, the instantaneous value of a random signal is a random value,

* Strictly, this occurs because this hypothetical experiment is not what is known as a stationary, ergodic process, and only with such processes can the principle be subject to classical spectral analysis. However, in practice no other alternative analysis method is available.

and has no specific significance. The properties which do have significance are parameters such as the mean, or mean-square value over all time and these by definition cannot be estimated from a short data duration. For example, some random event may perhaps occur on average, once in 1000 experiments, but if only 10, or even 100 experiments are observed, it will not be possible to estimate the occurrence rate accurately, however closely the results of the 10 or 100 observations are scrutinised.

There is one alternative to block-averaging to increase the statistical stability of computed spectra which is often advocated, and this is known as frequency smoothing. Essentially, instead of computing K spectra, and averaging the results at each frequency, the alternative is to average results at K adjacent frequency points. This approach offers no advantage in the trade-off between total data duration, and frequency resolution; if $N \times K$ samples are available, they can either be analysed as K blocks of size N , thus allowing a mean of K independent results, with a frequency spacing of S/N Hz (S is sample rate), or as one spectrum, with a spacing of S/NK Hz, but with the need to average K adjacent frequency points to obtain a similar stability. If frequency smoothing is used, the freedom from the requirement to segment the data for block averaging means that larger block sizes may possibly be used. This offers the advantage that each data point is seen in the context of more of its neighbouring points, thus 'making the most' of the information in the data in a way similar to overlap processing. On the other hand, because the computational time for the FFT of N points varies as $4N \log_2 N$, the fewer, longer, blocks will take longer to analyse, and will also need greater computer memory size.

3.5 Spectrum compensation

After adequate averaging, the computed spectrum is the best estimate of the signal spectrum as seen by the recording system. However, in some applications (for example when different designs or models of transducers are used for comparable data channels or when low order filters have been used, so that signals are attenuated at frequencies far below the Nyquist frequency) it may be desirable to modify the computed spectrum to compensate for the response characteristics of the transducer or signal conditioning. Provided the response characteristics are known, the compensation is basically straightforward, but the effect of compensation on quantization noise, and on aliased signals does deserve further brief discussion.

The process of spectrum compensation, and its effect on quantization noise is illustrated in Fig 16. A typical system response characteristic is shown in Fig 16a, while Fig 16b shows the inverse characteristic, *i.e.* the required compensation factor. As stated above, quantization noise (and also corrupt data noise) is white noise introduced downstream of the analogue part of the recording system. Thus when the compensation factor is applied to the white quantization noise, the resulting spectrum takes the form of the compensation factor, and the apparent spectrum of the quantization noise is now biased strongly towards high frequencies. When signal and quantization effects are combined, there will come some frequency beyond which the quantization noise will dominate the spectrum. This of course is a direct result of the fact that the response characteristics of the transducer, etc, progressively attenuate the signal

at higher frequencies and eventually a frequency will be reached where the signal is attenuated to below the level of the quantization noise. The only difference when spectrum compensation is applied is that instead of the signal falling off, the quantization noise grows stronger at the higher frequencies, and may appear to dominate the whole spectrum. As there is no point in studying data beyond the frequency where the signal has been attenuated below the quantization noise, it may be desirable to apply a band limited compensation factor, as shown in Fig 16c. The computed, compensated spectrum now resembles one obtained with the perfect low-pass filter shown in Fig 16d. However, in many applications, this approach will not be possible, because the frequency at which the quantization noise becomes dominant will depend on the strength of the recorded response, and this may vary from test to test. An alternative approach would be to calculate the expected quantization signal at each frequency, and to plot a 'quantization limit' on the computed spectrum. This simple expedient shows clearly when quantization effects are likely to be present, and was found quite useful in the VC 10 buffet research programme. Another alternative might be to subtract the expected level of quantization noise from the computed spectrum, but the random nature of the quantization noise makes this approach unattractive, particularly if (as in the case of the VC 10 buffet research programme) the data records were not sufficiently long for good statistical stability of truly random signals.

The effect of aliasing on spectrum compensation is shown in Fig 17. Because the compensation factor is applied to the data after any aliasing may have occurred, it applies the same compensation factor to all frequencies in the signal which alias to a given frequency in the sampled data. Thus it is equivalent to the 'multiple image' compensation factor, such as one of those shown in Fig 17a or 17b, applied to the unaliased signal. The difference between the two compensation factors shown is simply that one has been applied over the full frequency band up to the half sample rate, while the other applies compensation over a limited band only. The net result of these compensation factors and the basic system response characteristic, is shown in Fig 17c and 17d. It is immediately apparent that the price paid for a flat response at frequencies below the half sample rate is some amplification of signals above the half sample rate; *ie* of signals which alias. However, at any frequency in the computed spectrum, both 'true' signals, and signals which alias to the frequency are amplified by the same factor. Thus the compensation does not change the ratio of true to aliased signal at any position in the spectrum (indeed, it cannot, for by the time the stage of applying compensation is reached, there is no way to distinguish between the two). What it *does* do is to prevent the tendency to 'play down' the end of the spectrum where aliases will be least - well suppressed. The desirability (or otherwise) of the approach depends on the objectives of the analysis, and the nature of the data.

4 CROSS-SPECTRAL ANALYSIS

Cross-spectral analysis permits the relationship between two signals to be investigated in the frequency domain. Although in the past, cross-correlation techniques were favoured for such analysis, with the advent of the FFT, the direct computation of cross-spectral functions is now preferable. The relatively minor extensions to the FFT

algorithms which are required to compute the cross-spectral function will not be described here, but are readily available from any of the standard textbooks, eg Ref 1. Instead, the emphasis here will be directed towards presenting the significance and limitations of the cross-spectral functions.

Like the power spectrum described in the previous section, the cross-spectral functions are all continuous functions of frequency, and, with the FFT, are computed at equally spaced frequency intervals over the range from dc to the Nyquist frequency. In general, all the remarks in the preceding sections on quantization, sampling, aliasing, windows, block sizes, averaging and compensation etc, apply also to cross-spectral analysis, although in some cases the balance between the various conflicting requirements may be somewhat modified.

4.1 The cross spectrum

The basic 'building blocks' of cross-spectral analysis are the power spectra of the two signals, and the cross spectrum of the two signals. From these three basic functions, a number of further functions can be readily calculated. The power spectra have been described in the preceding section, and it was noted that the power spectrum describes the distribution of power within the frequency range, but it does not contain phase information. The phase information contained in the Fourier transform of the signal only relates the phases of the components of the signal to the instant at which the block of data starts, and this is not generally of any real significance. Thus if the Fourier transform of the signal is represented as a complex function of frequency (ie having real and imaginary components), then the power spectrum is equal to the square of the modulus of the transform at each frequency. By virtue of the properties of the conjugate of a complex number, the power spectrum can alternatively be expressed as the product of the transform, and its conjugate, at each frequency. In the standard notation of spectral analysis, this is expressed as

$$G_{XX}(f) = |X(f)|^2 = X(f)X^*(f)$$

where G_{XX} is the power spectrum of the signal X ,
 $X(f)$ is the Fourier transform of the signal X ,
and $*$ denotes the complex conjugate.

Similarly,

$$G_{YY} = |Y(f)|^2 = Y(f)Y^*(f) .$$

The cross spectrum between signal X and signal Y is now defined as:

$$G_{XY} = X(f)Y^*(f) .$$

Thus the cross spectrum is directly analogous to the power spectrum, except that it is computed from the transforms of two data channels. Indeed in the special case where the two signals are identical,

ie if

$$X(f) \equiv Y(f)$$

then

$$G_{XY} \equiv G_{XX} \equiv G_{YY} .$$

The one special point about the cross spectrum is that in general it is a complex function, ie it does contain information describing the phase relationship between the two signals at each frequency. The power level at each frequency in the cross spectrum is clearly equally dependent on the power levels of each of the input signals. The importance of averaging for power spectra has been explained in the preceding section. Because the cross-spectrum is a complex function, the effect of averaging is significantly different from the case of power spectra. When for some particular system, or at some particular frequency, the two signals are in some way directly related, the phase relationship between them will be constant. For example, in the single input, single output system shown in Fig 18a, the output signal Y occurs as a direct consequence of the input X , and the output is related directly to the input by the system transfer function, H . H may well be complex, but, at any given frequency, it is constant. Thus, whether X is deterministic or random, the phase and amplitude relationship between X and Y is constant. Thus the modulus of the averaged cross spectrum will be equal to the square root of the product of the averaged moduli of the individual cross spectra, and the phase angle of the averaged cross spectrum will be equal to the (constant) system phase relationship at the particular frequency. However, if there is not a direct relationship between the two channels at some frequency, and the output signal at that frequency is thus effectively a noise signal, the phase relationship will vary with time in a random manner. The analysis procedure will compute a phase angle for each block, but when the average is taken, the random variation in phase angle will make the complex power tend to zero, because effectively the real and imaginary components are averaged separately, and either may be positive or negative at random. Thus at frequencies where the two signals are strongly related, the averaged cross-spectrum shows a power level whose square is proportional to the product of the power in the two signals, and a phase angle which indicates the phase relationship between the signals. At frequencies where the two signals are not related at all, the averaged cross-spectrum will tend in the limit towards zero power, and a random phase angle. Of course, signals in the real world are seldom either totally related, or, at the other extreme, completely independent. Typical signals can be seen as a linear combination of the two extreme cases. They produce power in the cross-spectrum that is somewhat less than would be shown if the signals were totally directly related, and a phase angle that is an estimate of the phase difference between the related components in the signals. The cross spectrum is thus a very useful function, because it shows the phase relationship between the two signals, and the distribution of that power which is common to both signals, while, in the limit (ie with sufficient averaging) suppressing power which is not common to both signals.

Three further aspects of the cross-spectrum deserve mention. Firstly, although the phase information of the cross-spectrum describes the phase delay (or advance)

characteristics of the mechanism linking the signal, the power information in the cross-spectrum does not, by itself, relate to the power attenuation (or amplification) characteristics of the mechanism. (Techniques to study such characteristics will be described later.) Secondly, it is worth noting that it is quite possible for two signals both to have a peak at almost identical frequencies, but for the two peaks to be entirely unrelated. In such a case, the phase difference between the two signals would vary slowly, and given sufficient averaging time, the power in the cross-spectrum could be very low, despite the strong (but unrelated) peaks at the same frequency in both power spectra. Thirdly, because the power at any frequency measured in the cross-spectrum depends on three factors (the power in both power spectra, and the degree to which the phase angle is constant), it is not possible from the cross-spectrum alone to establish at which frequencies the phase angle is constant. High values of cross-spectral power may be due to the combination of a moderately constant phase relationship, and very high power levels in one or more signals.

4.2 The coherence function

This difficulty is overcome by use of the coherence function, which is obtained (at each frequency) by dividing the square of the modulus of the cross-spectrum, by the product of the two power spectra. This directly eliminates two of the three factors which cause the cross-spectrum to vary, and thus the coherence function is directly a measure of the degree to which the relationship between the two signals remains constant from block to block*. If at a particular frequency, the two channels are very strongly related, the coherence function will tend to the value 1, while if at some other frequency the two signals are entirely independent, the coherence function will have the value zero. Note that the coherence function is independent of the actual mean phase angle. The mathematical definition of the coherence function is

$$\gamma_{XY}^2(f) = \frac{|\overline{G_{XY}}(f)|^2}{\overline{G_{XX}}(f)\overline{G_{YY}}(f)} .$$

By careful scrutiny of the definition, it can be shown that two factors cause the coherence level to reduce from unity. It has already been shown that any variation of phase angle from block to block reduces the coherence, but it can also be shown that any variation in the ratio of the amplitude of the two signals at a particular frequency reduces the value of the coherence at that frequency. This second effect stems from the fact that the denominator in the coherence function relates to the product of the means of the two signals, whereas the numerator relates to the mean of a form of product function. Note that when two signals are directly related by some mechanism, the ratio of the amplitudes at any frequency (ie the gain) will be constant, thus again the coherence will be unity.

81014

* Thus the coherence function can be interpreted as indicating the variability of the phase relationship, and hence as a measure of the likely accuracy of the measured phase angles. This idea is developed further in Ref 21.

Because the coherence function effectively is a measure of the variation in the spectra from block to block, it must be remembered that no coherence information can be obtained from a single block of data in isolation. If the coherence is calculated for just one block, *ie* no averaging is implemented, a value of unity will be obtained at all frequencies, irrespective of the true value:

$$\gamma_{XY}^2(f)_{\text{no averaging}} = \frac{|G_{XY}(f)|^2}{G_{XX}(f)G_{YY}(f)} = \frac{G_{XY}(f)G_{XY}^*(f)}{G_{XX}(f)G_{YY}(f)} = \frac{X(f)Y^*(f)Y(f)X^*(f)}{X(f)X^*(f)Y(f)Y^*(f)} = 1 .$$

In fact it is known that the value of coherence will always be somewhat over-estimated unless a very large number of blocks is used to compute the average values of G_{XY} , G_{XX} and G_{YY} . The degree of bias depends on the true coherence level, and on the number of blocks used for averaging. No exact formula to compensate for this bias is known but Monte Carlo methods have been used by several authors to arrive at approximate correction factors. For example, Ref 22 suggests a correction of the form

$$\gamma_{\text{cor}}^2 = \gamma_{\text{raw}}^2 - \left(\frac{1 - \gamma_{\text{raw}}^2}{m} \right)$$

where m is the number of (non-overlapping) blocks used for each channel. However, the reference does not reveal which form of window function this result is appropriate to.

In the special case of input-output measurements on lightly-damped systems, another source of bias error can cause a reduction in computed coherence values. When the effects of finite data duration on power spectra were discussed in the previous section, it was noted that each spectral estimate represented not the power density at some particular frequency, but rather an average power density over some reasonably narrow frequency band. This is equally true of the cross spectrum. For a lightly-damped resonant system, the phase relationship between input and output changes very rapidly by 180 degrees between frequencies just below, and just above resonance. Thus an estimate of cross-spectral power at frequencies very close to resonance may contain contributions from frequencies above and below resonances. Because, unlike the power spectrum, the cross-spectrum is a complex quantity, the contributions differing in phase by 180 degrees will tend to cancel, thus reducing the cross-spectrum (but not the power spectra) at resonance. This can lead directly to a dip in input-output coherence at resonance for such systems. If meaningful input-output information is required in such circumstances, it is important that the effective resolution bandwidth is sufficiently narrow, so that only 'small' phase changes occur over the band. Similarly, it was noted earlier that the stored energy in this type of system causes the intensity of the response to vary less than the random variations in the intensity of the excitation (section 3.4). If such a system is analysed with a 'short' block duration (*ie* less than the time the system would take to stabilise at a steady amplitude of response, after a step change in the amplitude of a steady dynamic excitation), the ratio of the response amplitude to excitation amplitude will vary from block to block (because the response amplitude in one block will be significantly affected by the excitation amplitude of

preceding blocks). As a result, the measured coherence will be reduced. Because the frequency-spacing of the spectral estimates produced by the FFT is equal to the inverse of the block duration, this need for long blocks (from amplitude ratio considerations) corresponds directly to the need for close frequency spacing (from phase considerations).

When the signals are related directly by some linear mechanism at a particular frequency, the phase relationship, and the ratio of the amplitudes will stay constant. If the output signal is contaminated by noise, the noise will cause variations in both phase relationship and amplitude ratio, and will thus reduce the coherence. Hence for a single input, single output system with noise, as shown in Fig 18b, the coherence function can be taken as an indication of the signal-to-noise ratio at each frequency. Specifically, the relationship is:

$$\text{signal-to-noise} = \frac{\gamma_{XY}^2}{1 - \gamma_{XY}^2} .$$

Traditionally the coherence function has been used to study input-output relationships for the kind of simple system represented in Fig 18b. For systems that can be represented in that way, the coherence function will distinguish between the frequencies at which the output and input are related, and those at which they are not, *i.e.* at which the noise in the output measurements dominates any transmitted signal.

However, two cases can arise in practice which can cause the coherence function to overestimate the extent to which the output signal is a direct consequence of the input signal. The first case is illustrated in Fig 18c. Here, an unexpected feedback mechanism is permitting the output signal to influence the measured input. Because the feedback mechanism causes the input to be related to the output, the phase and amplitude ratios will stay constant from block to block, and the coherence will be high, even if there is no direct 'forward' path between input and output. The second case is shown in Fig 18d. Here some extraneous external signal is corrupting both input and output measurements. Again the amplitude and phase relationship between input and output remain constant (because both signals are dominated by the interference), and thus the coherence is high although there is no direct transmission path from input to output. It must be remembered that, because the coherence function represents the coherence of the system as a function of frequency, an actual coherence plot may be representative over part of the frequency band, but may contain one or both of the above deficiencies at other parts of the frequency band. A particularly common occurrence is for ac power supply frequencies to be present in both signals, and thus to produce spurious high coherence at those frequencies.

The author's experience of the use of the coherence function relates to its application to interpreting the vibrational response of a complex structure. In such applications, the coherence and phase relationships between accelerometers mounted around the structure can be used to determine the frequency of the dominant vibrations, and the corresponding mode shapes. This technique effectively exploits the fact that coherence measurements can identify 'extraneous' inputs which are common to two channels, since the two sensors are in effect both 'outputs' to a common external stimulus, in

contrast to the more usual application where transmission characteristics are investigated by analysis of 'input' and 'output' signals. The identification of modal frequencies, and the corresponding phase relationship for a pair of channels can be greatly facilitated by presenting the results of the cross-spectral analysis in the graphical format described in detail in section 5 below. To identify the overall response of a complex structure it is then necessary to study the results of such analysis for many pairs of channels, to build up a consolidated understanding. From this the phase relationships for the response throughout the structure can be identified, so that the mode shapes can be determined. In addition, the coherence measurements can distinguish between vibrations which are localised at a limited number of measuring stations, and others (sometimes at similar frequencies) which are significant throughout the structure. When describing the application of cross-spectral techniques to the identification of road vehicle vibrations, Refs 23 and 24 report that the significant frequencies could be identified automatically within the analysis software, by constructing histograms showing the number of coherence peaks occurring at each frequency within the band of interest. The close spacing of the important frequencies observed by the present author in the case of aircraft vibrations precluded such an automated procedure, but the technique does appear to have considerable merit when the response frequencies are reasonably well separated. In two other respects, the author's experiences of aircraft vibrations showed interesting parallels to the vehicle responses reported in Refs 23 and 24. Firstly, in both cases, strongly asymmetric vibrations were encountered, and secondly the phase relationships observed for both type of structure showed significant quadrature components in some modes, indicating a flow of power between distributed sources of excitation and damping.

Recently theoretical aspects of multiple-input, multiple-output systems have been the subject of much attention. This has led to substantial developments in the understanding and application of what are known as multiple, and partial coherence functions. When several, not necessarily independent, inputs drive a system, multiple- and partial-coherence techniques can be used to investigate to what extent the observed outputs are wholly due to a combination of the measured inputs, and also to what extent the observed response can be attributed to any individual input. These techniques were not applied to the VC10 buffet experiment, partly because they relate directly to the situation where the input forcing function is measured (the VC10 buffet programme was curtailed before measurements of any of the unsteady pressures which constitute the forcing function were made), and partly because of the extreme complexity of the method. The techniques will not be discussed further here, more details can be found in Refs 2, 4 and 25-30.

4.3 Non-symmetric cross-spectral functions

The definitions of the cross-spectrum, and also the coherence function are both 'symmetric' in that they both depend equally on the characteristics of the two signals that are being compared. It is this symmetry of dependence which permits the function to be regarded as describing what is common to the two signals, rather than describing what is transmitted from one channel to the other channel. However, when the physical nature of the system being studied justifies the interpretation that one signal is an

input, and the other signal is an output, several other, non-symmetric, spectral functions can be extremely valuable. The first of these is the transfer function, defined as:

$$H(f) = \frac{\overline{G_{YX}}(f)}{\overline{G_{XX}}(f)} .$$

Since the denominator is always real, the phase of the transfer function at each frequency is simply the same as that of the averaged cross-spectrum in the numerator. Note that the above definition of the transfer function gives a better estimate of the amplitude gain of the system than could be obtained by computing the square root of ratio of output power spectrum to the input power spectrum. This is because as noted previously the averaged cross-spectrum contains only signals that are common to both input and output (*i.e* that are coherent), and thus any noise contamination of the output signal is not mis-interpreted as having been transmitted through the system. For example, if the two channels are simply two independent noise sources, the coherence, and cross-spectrum, and hence transfer function between them will be zero, which would correctly represent the fact that there was no transmission mechanism between the two signals; nevertheless, the ratio of the two power spectra would be non-zero. One further point of interest is that, under the definition of transfer function given in the above equation, the product of the transfer function which would be calculated from X to Y , and that which would be calculated from Y to X , will in general be less than unity; because it will be equal to the coherence between X and Y ;

$$H_{X \rightarrow Y} H_{Y \rightarrow X} = \frac{\overline{G_{YX}} \overline{G_{XY}}}{\overline{G_{XX}} \overline{G_{YY}}} = \gamma_{XY}^2 = \gamma_{YX}^2 \leq 1$$

since

$$\overline{G_{XY}} = \overline{G_{YX}^*} .$$

Again, applied to the very simple example of two independent noise signals, the averaged cross spectrum would be zero, and thus so would the coherence, and both transfer functions.

Two further 'non-symmetric' spectral functions are the coherent output power, and the incoherent output power. These indicate the power at the output which is attributable to the input, and the power at the output which is independent of the input. The respective definitions are:

$$\begin{aligned} \text{coherent output power: } & \gamma_{XY}^2 \overline{G_{YY}} \\ \text{incoherent output power: } & (1 - \gamma_{XY}^2) \overline{G_{YY}} . \end{aligned}$$

An application of the coherent output power spectrum is described in Ref 31. These three 'non-symmetric' spectral functions were not used within the VC10 buffet analysis, because they are essentially based on an input and output to a modelled system, and are thus inappropriate in the absence of data on the forcing function. However, where the

appropriate information is available, they can be very powerful. There is also the possibility of extending the methods to multi-input systems.

4.4 Compensation of cross-spectral functions

The concept of filter compensation was discussed in the section on the power spectrum. Similar techniques can be applied to the cross-spectrum, and thus to the non-symmetric functions described above. Because the response characteristics of both signals appear both in the numerator and the denominator of the coherence function, compensation does not directly alter the spectral distribution of coherence at all*.

If the true phase relationship between the two signals is required accurately, it is important that the phase angles computed from the cross-spectrum are compensated for any difference in phase response of the sensors and signal conditioning used on the two channels. Such compensation was widely and successfully used in the VC 10 buffet programme. In addition, a further measure of phase compensation is required if the two channels are not sampled simultaneously. Synchronised sampling is widely, and erroneously, regarded as essential for the computation of cross-spectra and coherence functions. In practice, sequential sampling has substantial advantages in terms of hardware recording considerations. Simultaneous sampling can only be achieved by means of either one analogue-to-digital converter, fed by numerous sample and hold circuits, or alternatively by using numerous analogue-to-digital converters and, probably, digital stores to hold the values so that they can be recorded sequentially. Either method involves extra complexity, and may lead to degraded accuracy. Furthermore, many standard instrumentation systems (including the Plessey PV1513 system used for the present study) simply do not have the facility for simultaneous sampling. The most obvious result of non-simultaneous sampling is the appearance of a spurious phase lag between the two channels. The change of phase is proportional to the frequency, and also the interval between the samples of the two channels. Thus, the compensation to the measured phase relationship is straightforward. For example, if both channels have a common signal at frequency f Hz and the signal is in fact in phase, but one channel is sampled δt after the other, then there will be spurious phase difference of $360 \times f \times \delta t$ degrees between the channels. It is, of course, important that the time delay between the channels is constant; any jitter in the recording system will introduce random phase changes, and reduce coherence. A less obvious effect of sequential sampling is that the two data records will not in fact represent precisely the same time interval; one will start, and finish slightly before the other. If the signals are deterministic, this will be of no consequence, but if the signals are random, some specific event may occur while one channel is 'running', but after the other channel has stopped being recorded. This will lower the coherence. However, this effect is unlikely to be significant because of three factors. Firstly, the percentage of the data duration which is not common to both records will be very small; less than the time between two samples of one channel

* Indirectly, coherence measurements may be effected by the use of filters, because signal components which are attenuated will have poorer signal/noise ratios. However it is not possible to compensate for this effect unless the detailed spectral distributions of all noise sources (including phasing effects) is known.

compared to the time for typically the several hundred samples which make up the block. Secondly, as already stated, the coherence will only decrease if the data is random; the more deterministic the data, the less the coherence will be reduced. Thirdly, if any form of tapering data window is used, only a very small weighting will be applied to data at the very beginning and end of the block. Thus in practice no compensation to coherence will be required.

5 A CONVENIENT FORMAT FOR DISPLAYING PHASE AND COHERENCE INFORMATION

The cross-spectral analysis of VC 10 vibration data produced very many spectra, cross-spectra, and coherence plots. It became highly desirable to develop a convenient graphical means of displaying the information, so that the salient points could be absorbed readily. Inevitably the development of such a display was a compromise between suppressing irrelevant information, to avoid a cluttered format, and on the other hand, retaining as much useful information as possible. Although the format finally adopted was selected purely because it was found appropriate and useful for this particular application, it is described fully below for the benefit of any reader who may wish to apply it, or adapt it to suit the needs of similar cross-spectral analyses. In addition to the graphs and diagram described below, tabulations of both power spectra, cross-spectra, phase and coherence were produced for each pair of channels analysed.

The graphical output comprised plots of power spectra, modulus of cross-spectrum, and coherence function, and in addition a combined phase-coherence-power diagram. All except this last mentioned diagram are completely conventional, and do not require further explanation. The phase-coherence-power diagram combined these three key parameters describing the relationship between the two signals. An example of the diagram, which is a perspective representation of a three-dimensional plot in cylindrical coordinates, is shown in Fig 19. The locus of the value of coherence (plotted radially, on a linear scale) and phase angle, as frequency is made to vary are plotted on a 'ground plane'. Each frequency point of the FFT is marked with a cross, and a process of interpolation by equiangular spirals is used to joint adjacent frequency points. The nature of the data causes the plotted locus to take the form of a series of 'lobes', each related to a mode of vibration which is coherent between the two subject signals. Simple software logic (described in detail below) recognises the existence of the lobes, and causes the cross-spectral power at the dominant frequency to be represented as a line in the axial direction. A logarithmic scale is used for the power axis, as is usually the practice in spectral analysis. In addition, the frequency of the mode is annotated beside the power line. Thus by studying the phase-coherence locus on the ground-plan, the presence of coherent vibrations could be seen, and the associated phase angles identified. Then by means of the cross spectrum (axial) scale, the relative strengths of the different modes of vibration could be assessed, and also the corresponding frequencies could be read off.

The system finally adopted to distinguish between frequencies of importance (where the cross-spectral power, and frequency are indicated), and unimportant frequencies is described below. It must be stressed that the system was adopted because it was simple, and yet gave useful results.

The system was as follows:

- (a) Points were never flagged as being significant if the coherence was less than 50%.
- (b) Points were only flagged if they corresponded to a peak (ie had a value greater than either of the adjacent frequency points) in either of the power spectra, or in both coherence and cross-spectrum plots.

Very occasionally this system caused a point to be ignored when an assessment of the data by eye might indicate that it could be flagged; this was almost always due to the point failing to reach the 50% coherence level because it was part of a data set with low signal levels, and hence poor signal-to-noise ratio and low coherence values at all frequencies. This could possibly have been improved by scheduling the coherence threshold in relation to the overall coherence level of the data. More frequently, the system allowed more than one point on each lobe to be identified. This could happen if the true response frequency fell between two of the FFT frequency points. In this case, the power at the points either side of the true frequency could be very similar, and small random variations could cause one to be a peak in the power spectrum of the first channel, and the other point to be a peak in the power spectrum of the second channel.

Because in almost all cases, the coherence tended to be low at high frequencies, the data was scanned to identify the frequency above which coherence never exceeded 50%. To reduce clutter in the diagram, the phase/coherence locus was not plotted beyond this frequency.

6 CONCLUDING REMARKS

The methods and procedures outlined above have been condensed from information reported in a wide range of sources, and have been refined through practical experience. The power and usefulness of the techniques was well displayed in the successful interpretation of the complex vibrations measured on an aircraft in buffet. Accordingly, a substantial amount of the material presented is of special relevance to such uses as the identification of complex vibrations, based on limited data. However, the more general topics of quantization, aliasing, windowing, practical application of the FFT, and the averaging of spectra are discussed in full, as are the more specialised topics of cross-spectral analysis and the coherence function. As a practical guide, the report is thus intended to be of value to anyone about to embark on spectral or cross-spectral analysis in whatever field of application.

Appendix ATHE CONVOLUTION THEOREM

According to the convolution theorem, when two signals are multiplied together to generate a third signal, the Fourier transform of this third signal may be calculated by performing the convolution operation between the Fourier transforms of the two original signals. The convolution operation is denoted by the * sign, and is defined as follows:

$$A(f) * B(f) = C(f) = \int_{x=-\infty}^{x=+\infty} A(x)B(f-x)dx .$$

The convolution operation can be visualised as follows.

To evaluate the result of the convolution, $C(f)$ at some particular value of $f = f'$ imagine the two functions $A(f)$ and $B(f)$ plotted out, but with $B(f)$ plotted with the origin offset by f' and with the f -axis reversed (Fig A1). Now construct the graph of the product of the two functions. The value of the convolved function, at the point $f = f'$ is now equal to the integral of this product (limits $-\infty$ to $+\infty$). Note that the functions A , B and C will in general be complex. To evaluate $C(f)$, for all f in this way is obviously a complicated procedure, and thus the method is only of value as a practical procedure if either $A(f)$ or $B(f)$ are such as to make the integration very straightforward. Fortunately this is in fact the case for several important facets of spectral analysis.

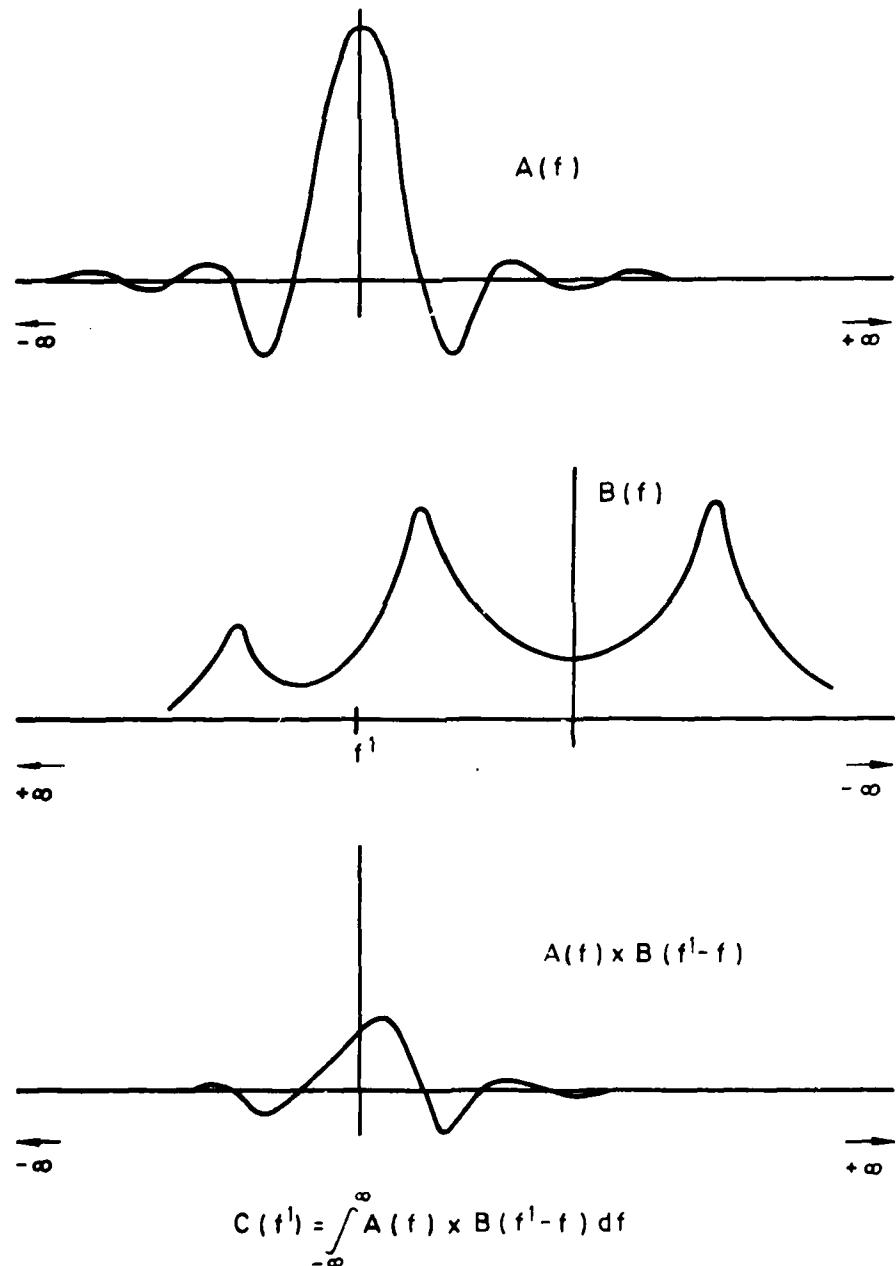


Fig A1 The convolution process

Appendix B

DEMONSTRATION OF THE SPECTRAL CONSEQUENCES OF SAMPLING

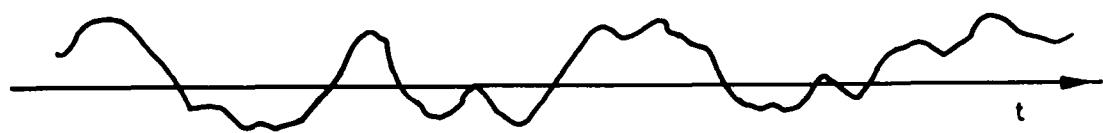
Sampling may be regarded as multiplying a continuous signal by a sampling function which consists of a regular train of sample pulses. Ideally each pulse has unit energy and infinitesimal duration. Such idealised pulses are generally referred to as delta functions*. (The effects of non-idealised sampling pulses will be discussed later.) To find the Fourier transform, and hence the spectrum of the sampled signal, the transform of the continuous signal can be convolved with the transform of the sampling function. The process is illustrated in Fig B1. The transform of the sampling function is in fact a series of real delta functions at 0 Hz and at all harmonics of the sample rate. When carrying out the integration in the convolution process, the transform of the sampling function may be treated as the sum of a series of delta functions, each to be convolved and integrated separately, and then the results summed to produce the final answer.

The application of the convolution process may conveniently be illustrated by considering firstly the component of the transform of the sampling function which is a delta function at 0 Hz. In this case, application of the procedure of convolution as described in Appendix A produces a transform which is identical to that of the original (continuous) data. Viewed in the time domain, the delta function at 0 Hz represents a steady 'dc' level, which, when multiplied with some signal, will not change the transform of that signal (Fig B2). When the convolution process is performed with each of the other delta functions in the transform of the sampling function, the result is similar, except that the resulting transform is shifted in frequency by an amount equal to the frequency of the particular delta function. The resulting power spectrum, when the contributions of all the delta functions are added together, is a series of regularly-spaced images of the original spectrum, as shown in Fig B1(f). In the example shown, the sample rate is sufficiently high, so that adjacent images of the signal spectrum do not overlap when the data is sampled. However, it can readily be seen that a lower sample rate would cause the peaks of the sampling function spectrum to move closer together, and the spectrum of the sampled signal could become impossible to interpret (Fig 5).

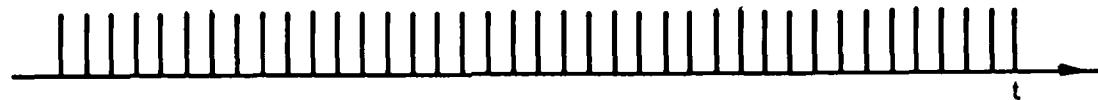
If the peaks in the sampling function are of finite width, the effect on the spectrum of the function is merely to reduce the amplitude of the higher frequency peaks. However, as it is in general the lower-frequency peaks which cause the most severe

* Strictly, the sampled data is a function which is *undefined* except at the sample instants, whereas the product of a continuous signal, and a train of delta functions, is a function which is *zero* except at the sample instants. It is nevertheless valid to calculate the spectrum of such a function, and to use it as a description of the *effective spectrum* of the sampled data, because any analysis procedure applied to the sampled data represents continuous integrals over time by summations at the sample instants. This representation would only be exact in the case of a signal which was zero except at the sample instants. Thus the procedure used to calculate the spectrum corresponds exactly to the way in which the sampled data is actually interpreted by any form of analysis. This is true even if the analysis procedure actually represents the data by some scheme of interpolation between the sample values, because this interpolation is in itself a procedure which represents continuous input data by a sequence of samples. The only difference in this case is that the interpolation process is effectively a filter, whose characteristics should be superimposed on the spectrum of the sampled data.

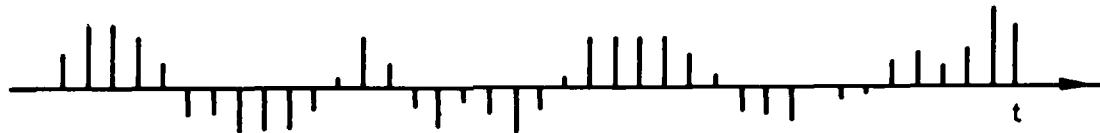
aliasing problems, the effect of sampling function pulse shape (ie how good the analogue-to-digital conversion is at taking an instantaneous value of the signal, rather than an average over a short time interval) on aliasing is usually negligible.



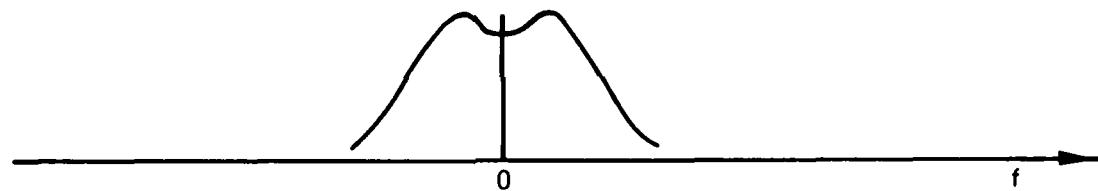
a Continuous signal



b Sample pulses (S samples/second)



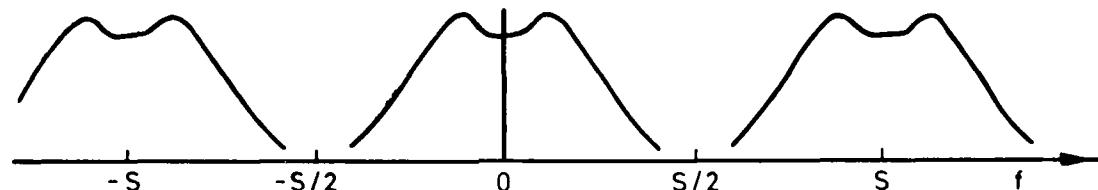
c Sampled values = product of continuous signal and sample pulses



d Spectrum of continuous data (symmetric about f=0 since signal is real)



e Spectrum of sample pulses (delta functions at f=0, ±S, ±2S, etc)



f Spectrum of sampled data - aliased image spectra arise from convolution of transform of signal with transform of sample pulses

Fig B1a-f Derivation of the spectrum of sampled data

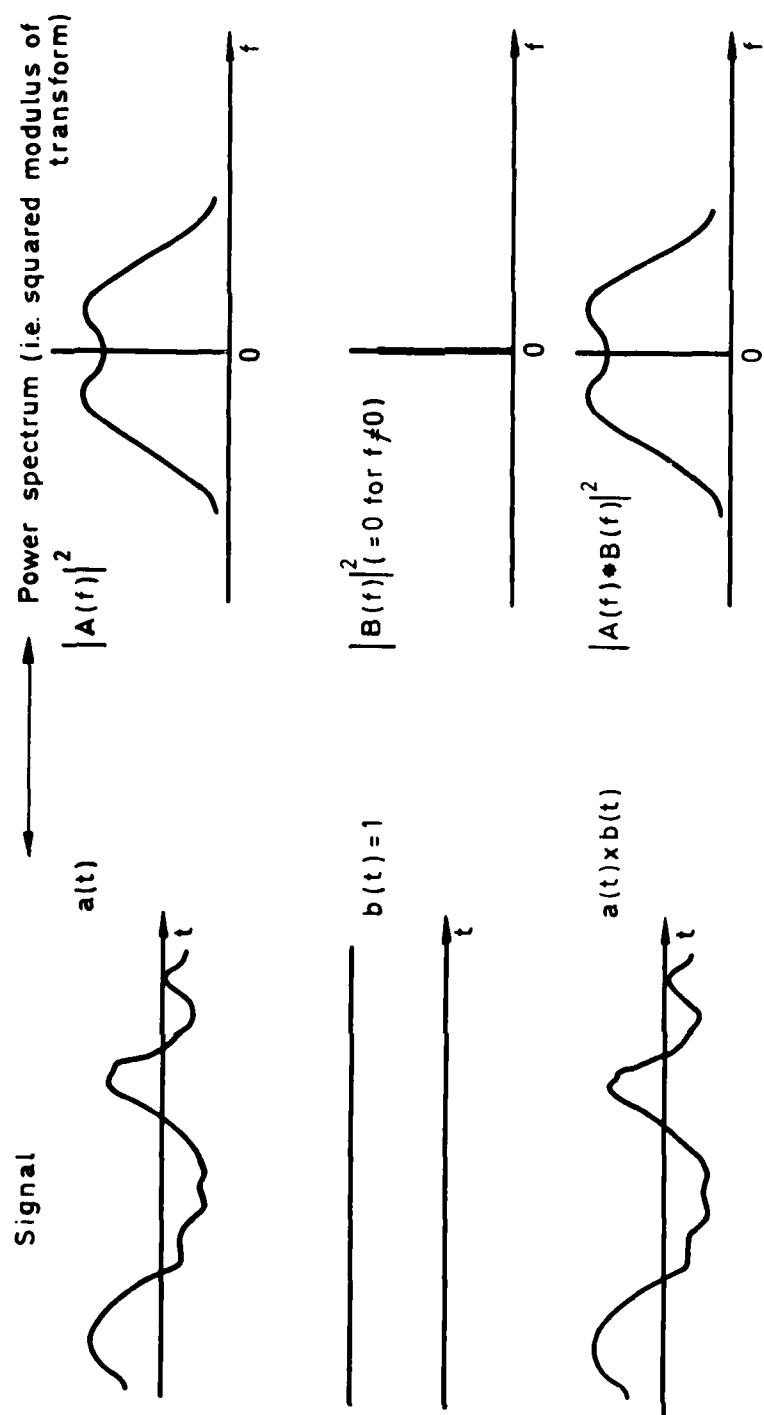


Fig B2 Time-domain multiplication and frequency domain convolution – a simple case

Appendix C

MATHEMATICAL DEFINITION OF FOURIER TRANSFORM AND POWER SPECTRUM

If $x(t)$ is a function of time ('the signal') then its Fourier transform, $X(f)$ is defined as

$$X(f) = \int_{t=-\infty}^{t=+\infty} x(t) e^{-j2\pi ft} dt , \quad (C-1)$$

alternatively, $x(t)$ may be expressed as the inverse Fourier transform of $X(f)$:

$$x(t) = \int_{f=-\infty}^{f=+\infty} X(f) e^{j2\pi ft} df . \quad (C-2)$$

The power spectral density function $G_{XX}(f)$ is defined as

$$G_{XX}(f) = |X(f)|^2 = X(f)X^*(f) . \quad (C-3)$$

Note that equation (C-1), and hence (C-3), require that the function $x(t)$ be known for all time. However, in the practical case, a spectrum computed from data restricted to a finite duration can be taken as an estimate of the true spectrum. The estimate will be representative of the spectrum over all time provided that the spectral composition of the signal does not vary systematically with time. A more accurate estimate can be obtained by averaging the power spectrum computed from numerous independent data records, as explained in section 3.4.

LIST OF SYMBOLS

$G_{XX}(f), G_{YY}(f)$	power spectra of signals $x(t), y(t)$
$G_{XY}(f)$	cross-spectrum of signal $x(t)$ and $y(t)$
f	frequency (Hz)
$H(f)$	transfer function
ℓ	number of frequency points in discrete spectrum
m	number of spectra to be averaged
N	number of samples in each 'block'
r	number of zeros used in zero padding
s	sample rate (second ⁻¹)
t	time (seconds)
T_B	block duration (seconds)
$x(t), y(t)$	functions of time ('signals')
$X(f), Y(f)$	Fourier transform of $x(t)$ and $y(t)$
X_i	i th point of discrete Fourier transform
Z	factor relating minimum frequency to sample rate in zoom transform (Hz)
$\gamma_{XY}^2(f)$	coherence function of signals $x(t)$ and $y(t)$
δf	frequency spacing of spectral estimates in discrete Fourier transform
δt	sampling delay between two signals (seconds)
ϵ	normalised standard error
τ	period or duration (seconds)
$\overline{\quad}$	mean of many independent spectra
$ \quad $	modulus
$*$	complex conjugate or convolution operator

Note that for convenience, the dependence of many of the above parameters on frequency is dropped in places in the text, eg $H(f)$ may be written simply as H .

REFERENCES

<u>No.</u>	<u>Author</u>	<u>Title, etc</u>
1	J.S. Bendat A.G. Piersol	<i>Random data: Analysis and measurement procedures.</i> John Wiley & Sons, New York (1971)
2	J.S. Bendat A.G. Piersol	<i>Engineering applications of correlation and spectral analysis.</i> John Wiley & Sons, New York (1980)
3	R.K. Otnes L.D. Enochson	<i>Digital time series analysis.</i> John Wiley & Sons, New York (1972)
4	R.K. Otnes L.D. Enochson	<i>Applied time series analysis. Vol.1.</i> John Wiley & Sons, New York (1978)
5	S.L. Buckingham	RAE Technical Report (to be published)
6	L.W. Gardenhire	Sampling and filtering. In <i>Basic principles of flight test instrumentation</i> , edited by A. Pool and D. Bosman, AGARDograph No.160, Vol.1 (1974)
7	A.M. Furber	The interpretation of discrete Fourier transforms. RAE Technical Report 73127 (1973)
8	J.C. Burgess	On digital spectrum analysis of periodic signals. <i>J. Acoust. Soc. Am.</i> , 58, Part 3, 556-567 (1975)
9	C. Bingham M.D. Godfrey J.W. Tukey	Modern techniques of power spectrum estimation. <i>IEEE Trans Audio Electroacoust</i> , AU-17, Part 2, 56-66 (1967)
10	A.W.R. Gilchrist	The use and properties of window functions in spectrum analysis. CRC (Canada), 1221 (1971)
11	T.S. Durrani J.M. Nightingale	Data windows for digital spectral analysis. <i>Proc. IEE</i> , 119, Part 3, 342-351 (1972)
12	A. Papoulis	Minimum-bias windows for high-resolution spectral estimates. <i>IEEE Trans Inf. Theory</i> , IT-19, Part 1, 9-12 (1973)
13	A. Eberhard	An optimal discrete window for the calculation of power spectra. <i>IEEE Trans Audio Electroacoust</i> , AU-21, Part 1, 37-43 (1973)
14	K.M.M. Prabhu J.P. Agrawal V.U. Reddy	Performance comparison of data windows - digital filtering and prediction theory. <i>Electron Lett</i> , 13, 600-601 (1977)
15	K.M.M. Prabhu V.U. Reddy	Data windows in digital signal processing - a review. <i>J. Inst. Electron. Telecommun. Engrs.</i> , 26, 69-76 (1980)
16	K.M.M. Prabhu J.P. Agrawal	Selection of data windows for digital signal processing. In <i>Proceedings of the 1978 IEEE International Conference on Acoustics, Speech and Signal Processing</i> , pp 79-82 (1978)

REFERENCES (continued)

<u>No.</u>	<u>Author</u>	<u>Title, etc</u>
17	K.M.M. Prabhu J.P. Agrawal	Time and frequency domain implementations of data windows. <i>Archiv Fuer Electronik und Uebertragungstechnik</i> , 33, 224-226 (1979)
18	T.G. Brown C.G. Brown J.C. Hardin	Program for the analysis of time series. NASA TM X-2988 (1974)
19	J.B. Roberts D. Surry	Some experiments with 'on-line' spectral analysis using a small digital computer. NPL (Aero) Report 1328 (1971)
20	G.C. Carter C.H. Knapp A.H. Nuttall	Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. <i>IEEE Trans. Audio Electroacoust</i> , AU-21, Part 4, 337-344 (1973)
21	C.R.S. Talbot	Coherence function effects on phase difference determination. <i>J. Sound Vibration</i> , 39, Part 3, 345-358 (1975)
22	V.A. Benignus	Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform. <i>IEEE Trans. Audio Electroacoust</i> , AU-17, Part 2, 145-150 (1969)
23	C.R.S. Talbot G.H. Tidbury	Determination of vibrational modes of vehicle structures by the application of random process theory. In <i>Proceedings of the Society of Environmental Engineers Conference SEEKO 73</i> (1973)
24	C.R.S. Talbot G.H. Tidbury S.K. Jha	Identification of the vibrational modes of a car driven on the road. In <i>Proceedings of ASME Conference on design engineering</i> , 75-DET-115 (1975)
25	L.D. Enochson	Frequency response functions and coherence functions for multiple input linear systems. NASA CR-32 (1964)
26	C.J. Dodds J.D. Robson	Partial coherence in multivariate random processes. <i>J. Sound Vibration</i> , 42, Part 3, 243-249 (1975)
27	J.S. Bendat	Solutions for the multiple input/output problem. <i>J. Sound Vibration</i> , 44, Part 3, 311-325 (1976)
28	J.S. Bendat	System identification from multiple input/output data. <i>J. Sound Vibration</i> , 49, Part 3, 293-308 (1976)
29	R. Potter	Matrix formulation of multiple and partial coherence. <i>J. Acoust. Soc. Am.</i> , 61, Part 3, 776-781 (1977)

REFERENCES (concluded)

<u>No.</u>	<u>Author</u>	<u>Title, etc</u>
30	S. Barrett R.M. Halvorsen	The use of coherence functions to determine dynamic excitation sources on launch vehicle payloads. NASA CR 3142 (1979)
31	W.G. Halvorsen J.S. Bendat	Noise source identification using coherent output power spectra. <i>J. Sound and Vibration</i> , 15-24, August 1975


A handwritten signature in black ink, appearing to read "W. G. Halvorsen".

Fig 1a-d & Fig 2a-d

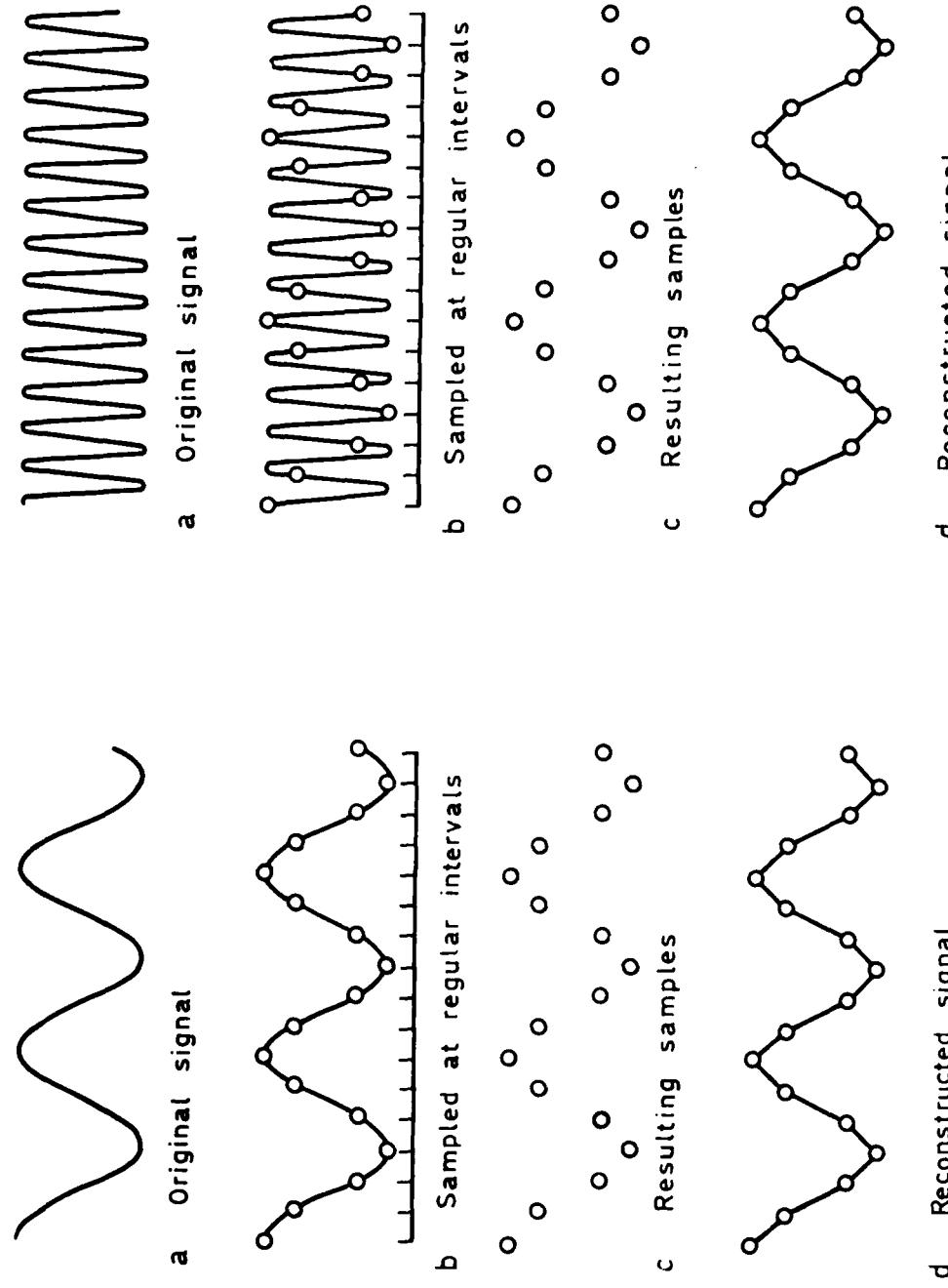
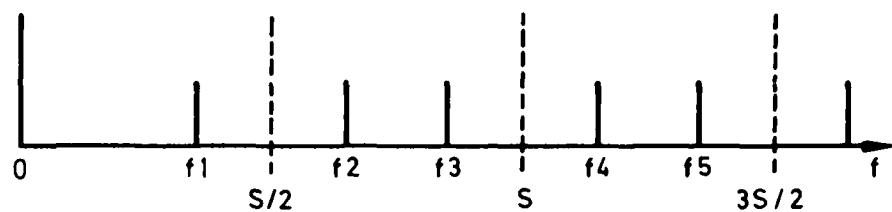


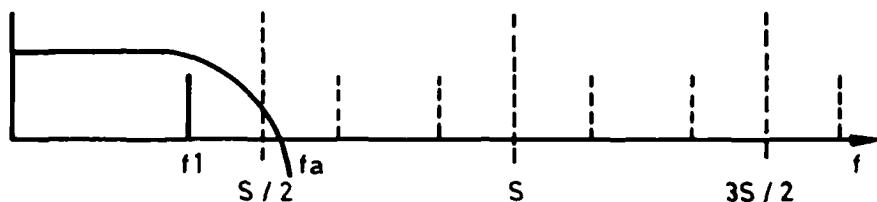
Fig 1a-d Sampling and reconstruction:—
adequate sample rate

Fig 2a-d Sampling and reconstruction:—
inadequate sample rate

Fig 3a&b & Fig 4a-c

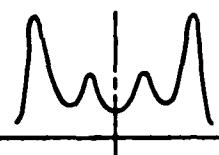


a Any linear combination of signals at frequencies f_1, f_2, \dots could give rise to the same observed samples

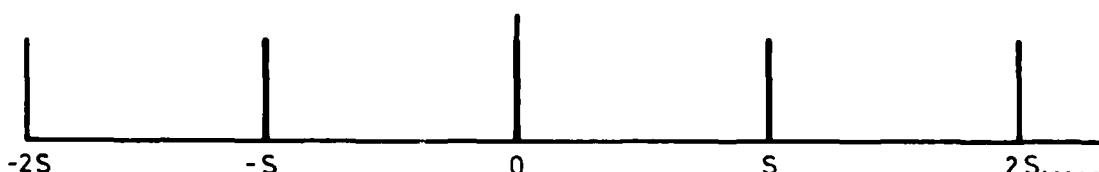


b But if it is known that the input signal was restricted to the band 0- f_a , the actual signal can only have been at f_1

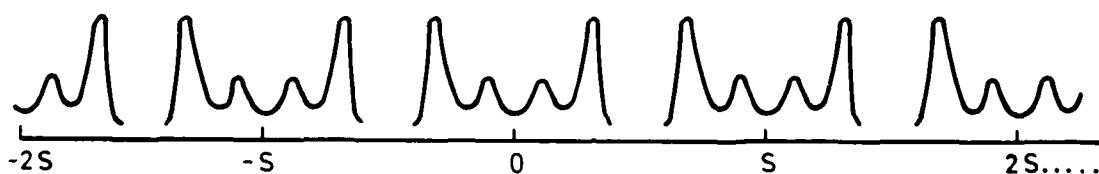
Fig 3a&b



a Spectrum of continuous data

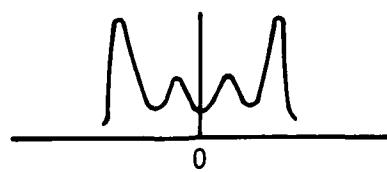


b Spectrum of sampling function

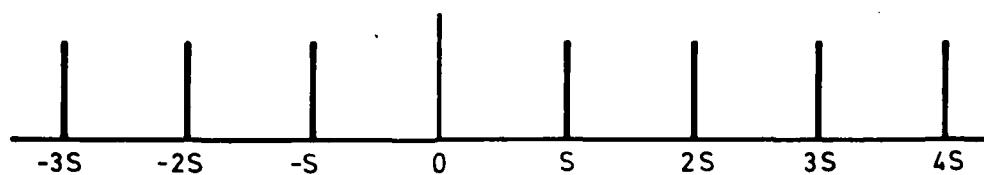


c Spectrum of sampled data

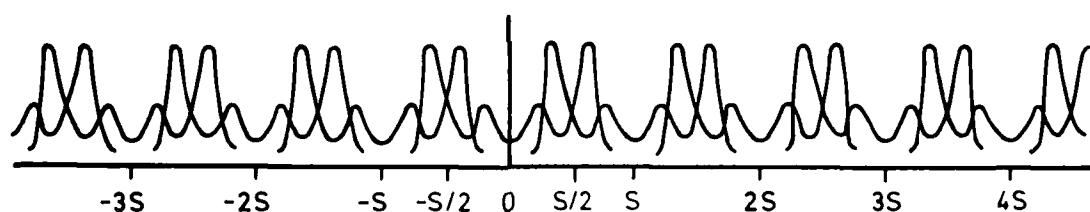
Fig 5a-d



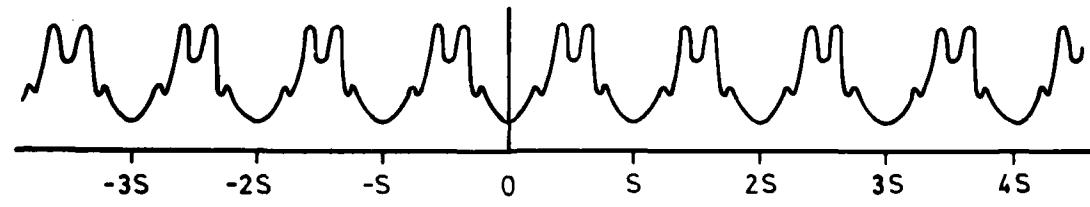
a Spectrum of continuous signal



b Spectrum of sampling function (sample rate lower than shown in Fig 4)

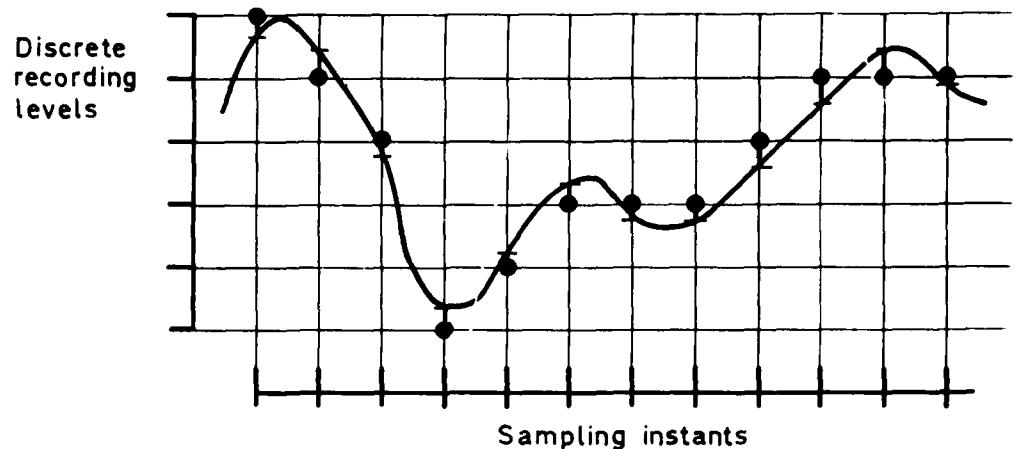


c Spectrum of sampled data - shown as overlapping images



d Spectrum of sampled data when contributions from overlapping images are combined

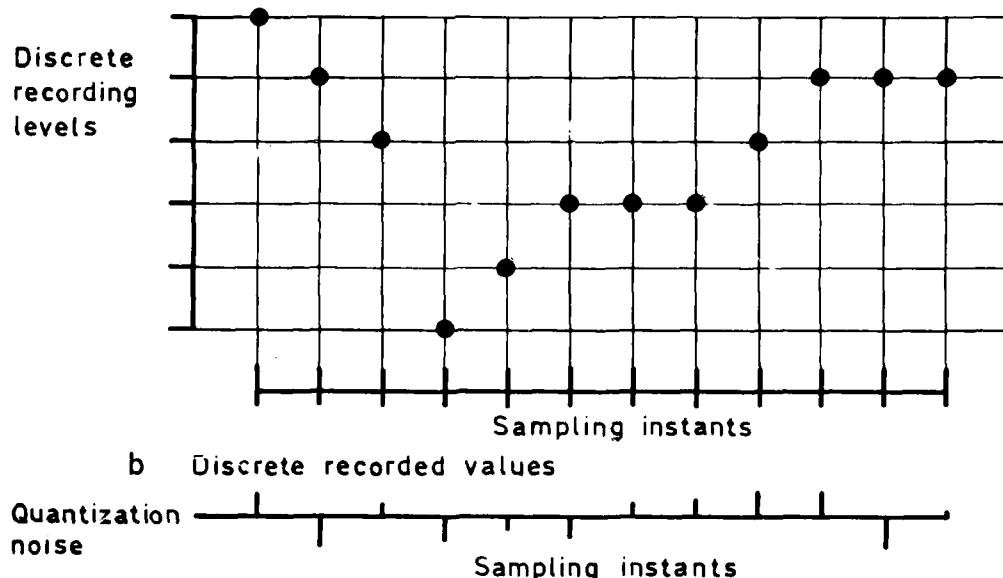
Fig 5a-d Inadequate sample rate causes images to merge



Key:

- True continuous signal
- True value at sampling instant
- Closest discrete digital value

a Comparison of continuous signal and discrete recorded values



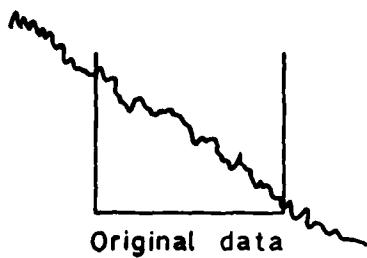
b Discrete recorded values

Quantization noise

Sampling instants

c Quantization noise is difference between true value at sample instants and discrete recorded values

Fig 7



When a block of data contaminated by drift is analysed, the drift is misrepresented as the sum of a series of components which are low-order harmonies of the block rate:-



These spurious components can be eliminated by fitting a straight line to the data within the block (using a least-squares fit), and then analysing the residual values after this drift-line has been subtracted from the data

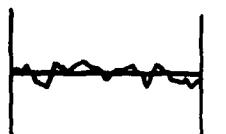
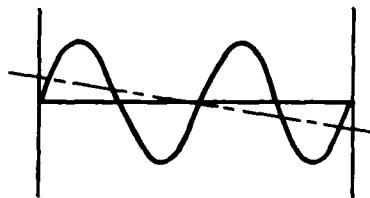
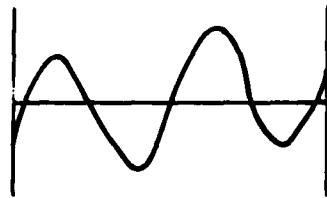


Fig 7 Elimination of drift

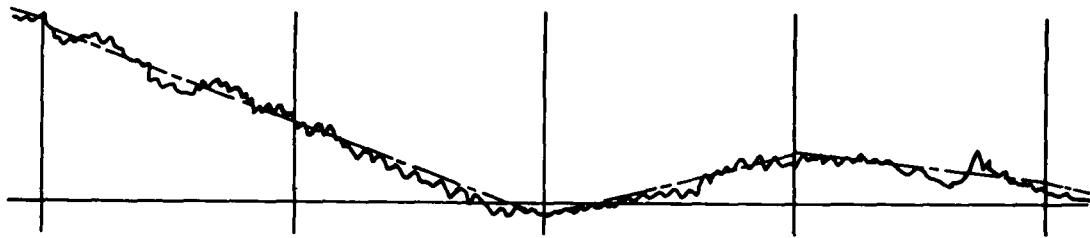


a Simple signal with no drift, and drift line calculated by least-squares fit

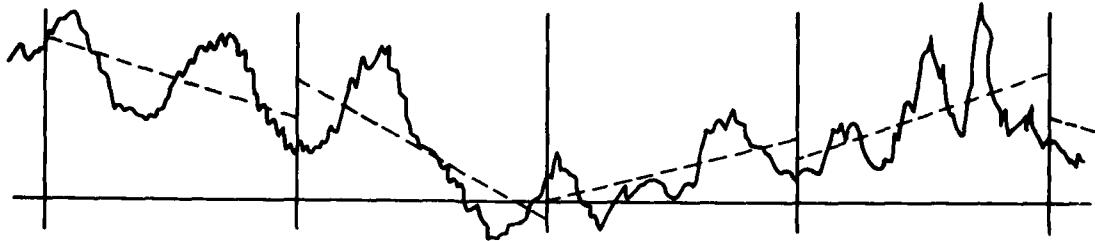


b Subtraction of apparent drift actually introduces a drift, and distorts true signal

Fig 8a&b When no drift is present, drift elimination may distort signal



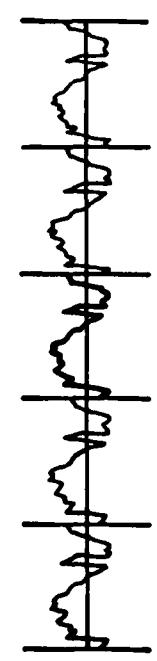
a When drift is genuine, the discontinuity of the drift lines at the block boundaries tends to be small



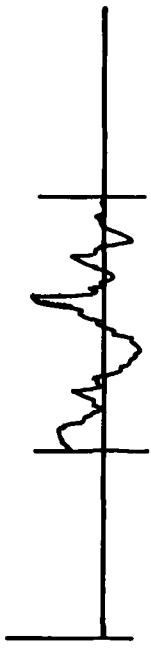
b Apparent drift caused by phasing effects can show quite large discontinuities at the block boundaries

Fig 9a&b Identification of genuine drifts

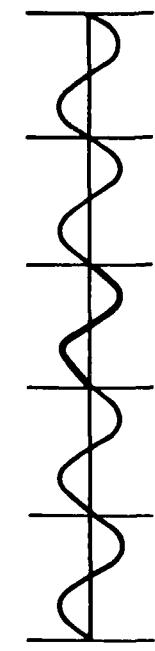
Fig 10a-c & Fig 11a-c



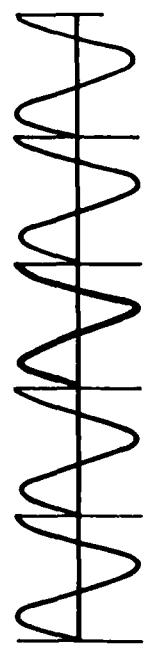
a In analysis, each block of data is treated as though it repeats cyclically for all time



a Data record of finite duration =



b Components of the signal which are exactly periodic in the block duration are unaffected



c Components of the signal which are not exactly periodic in the block duration are interpreted as having transients at the block boundaries



b Signal of unlimited duration \times

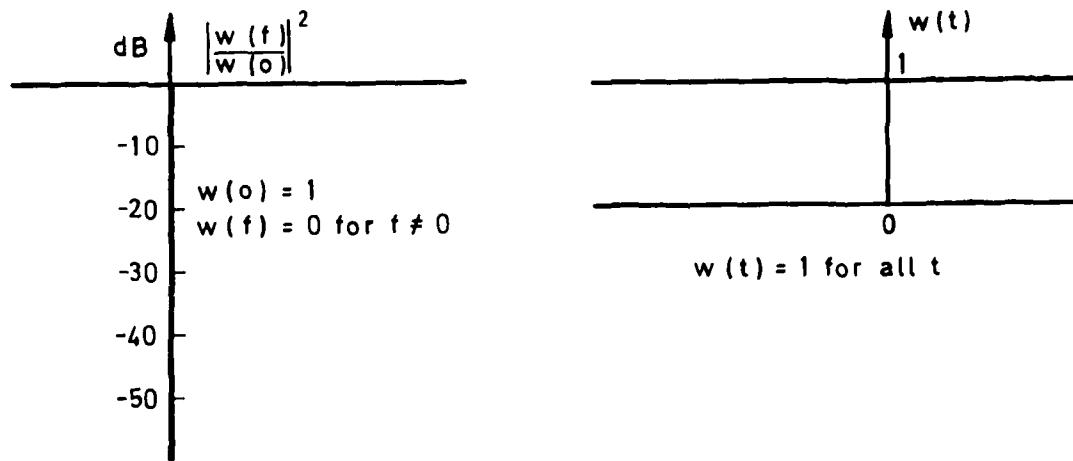


c Duration-limiting window function

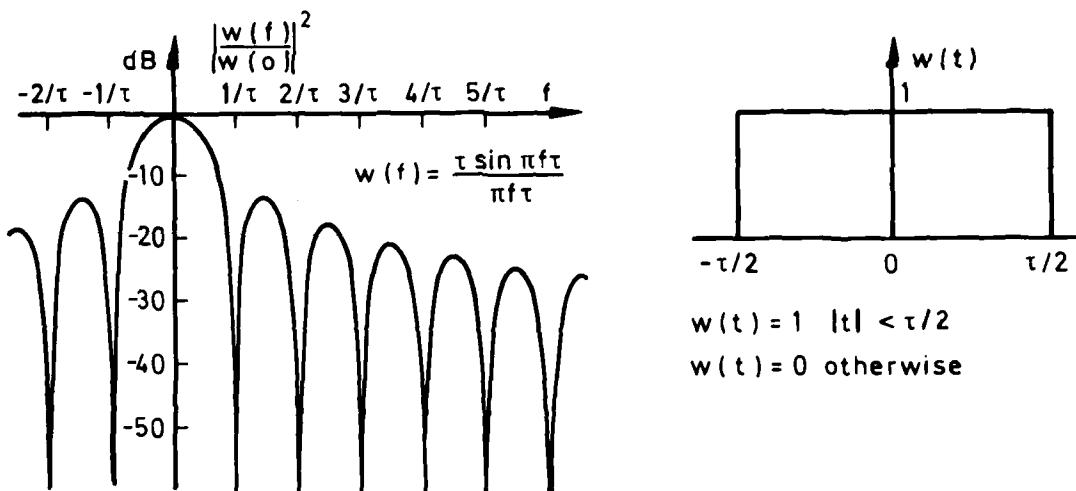
Fig 10a-c Distortion of signals non-periodic in block duration

Fig 11a-c Application of the window function

Fig 12a&b

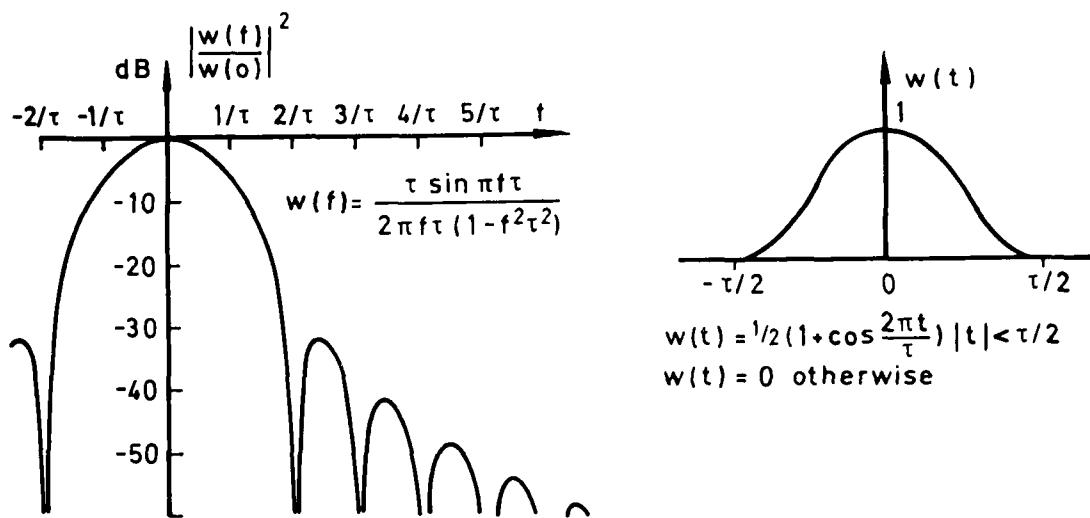


a Perfect (non-blurring) window - impossible to realize

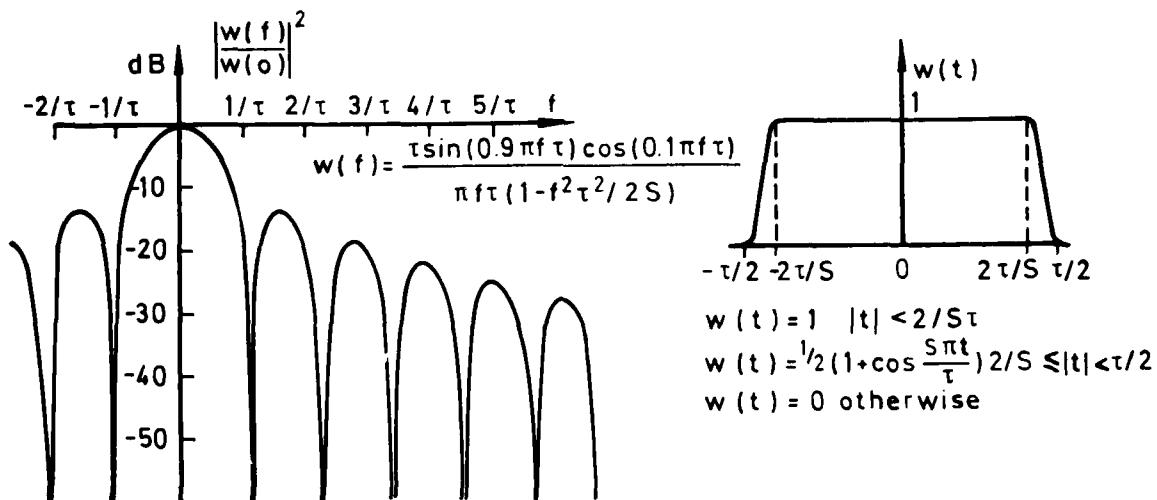


b Box car window

Fig 12c&d



c Hann of cosine bell window



d 90% cosine taper window

Fig 12e&d Window functions in the frequency and time domains

Fig 13 & Fig 14a-e

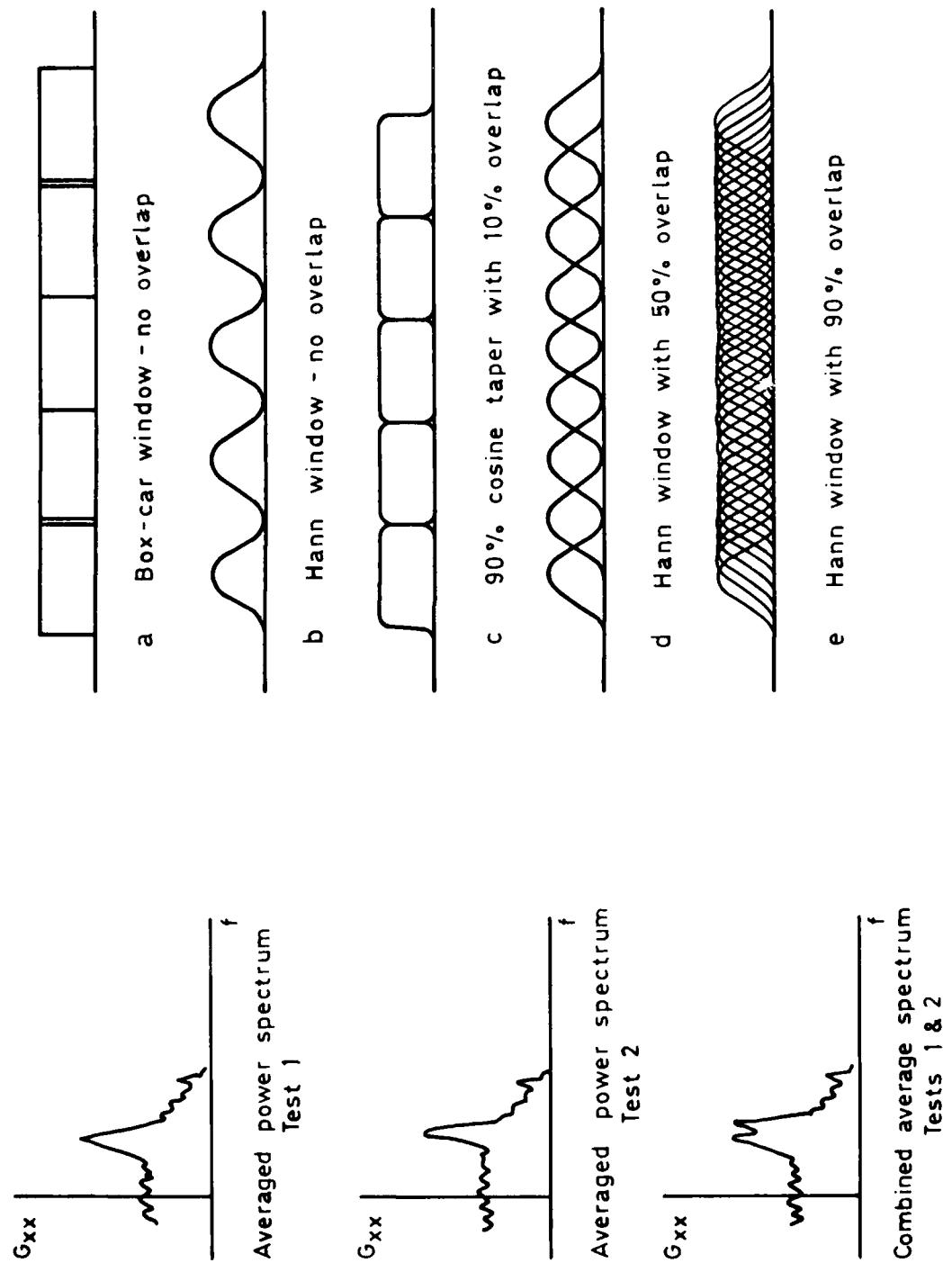


Fig 13 Averaging of independent tests

Fig 14a-e Windowing and overlapping

Fig 15a-c & Fig 16a-d

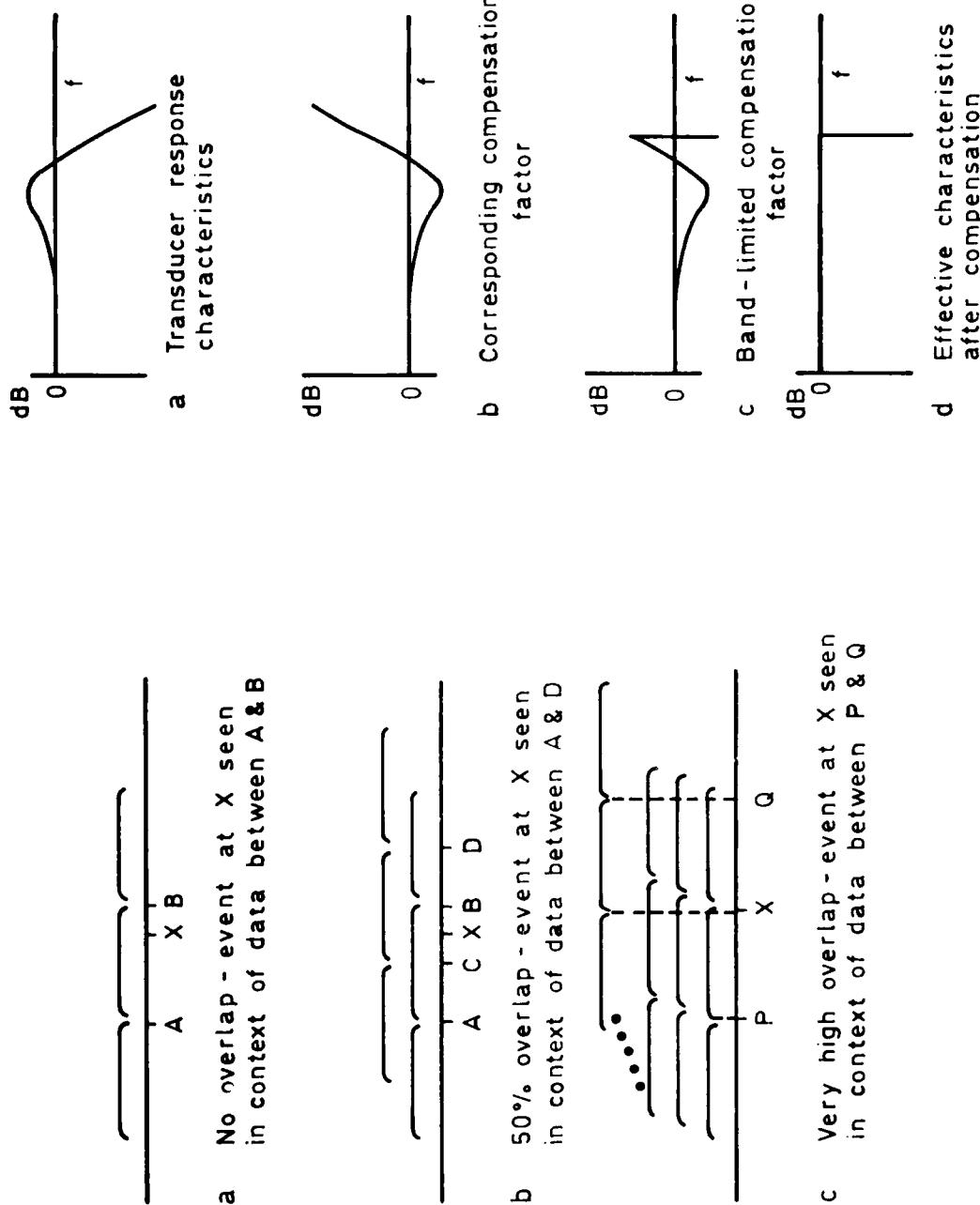


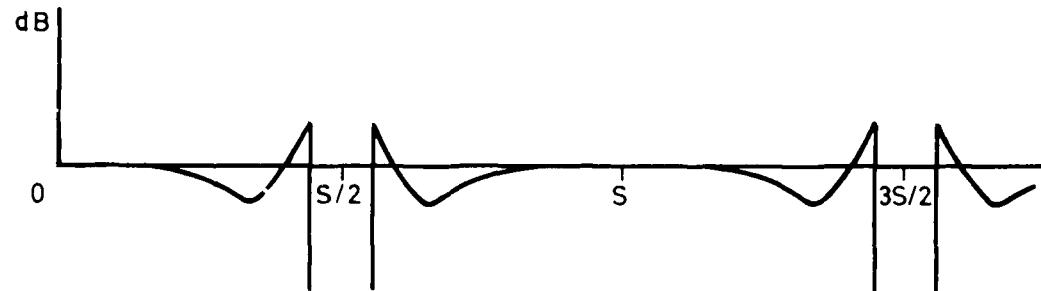
Fig 15a-c Benefits of overlapping

Fig 16a-d Spectrum compensation

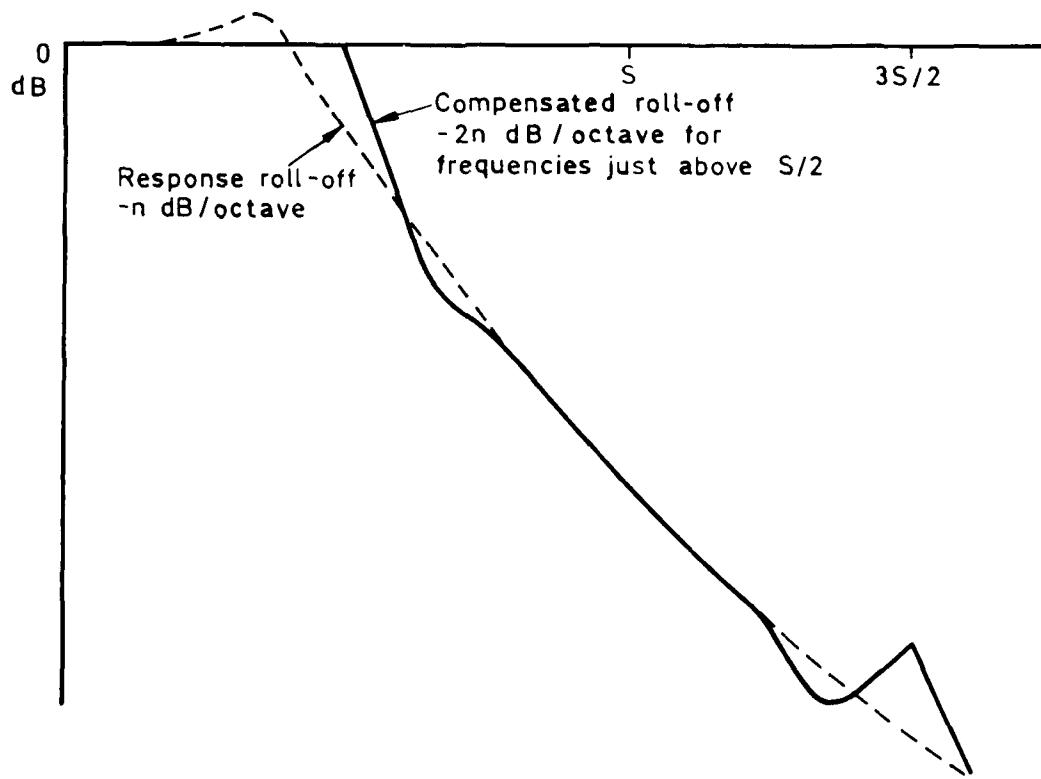
Fig 17a-c



a Effective compensation factor - full - band compensation

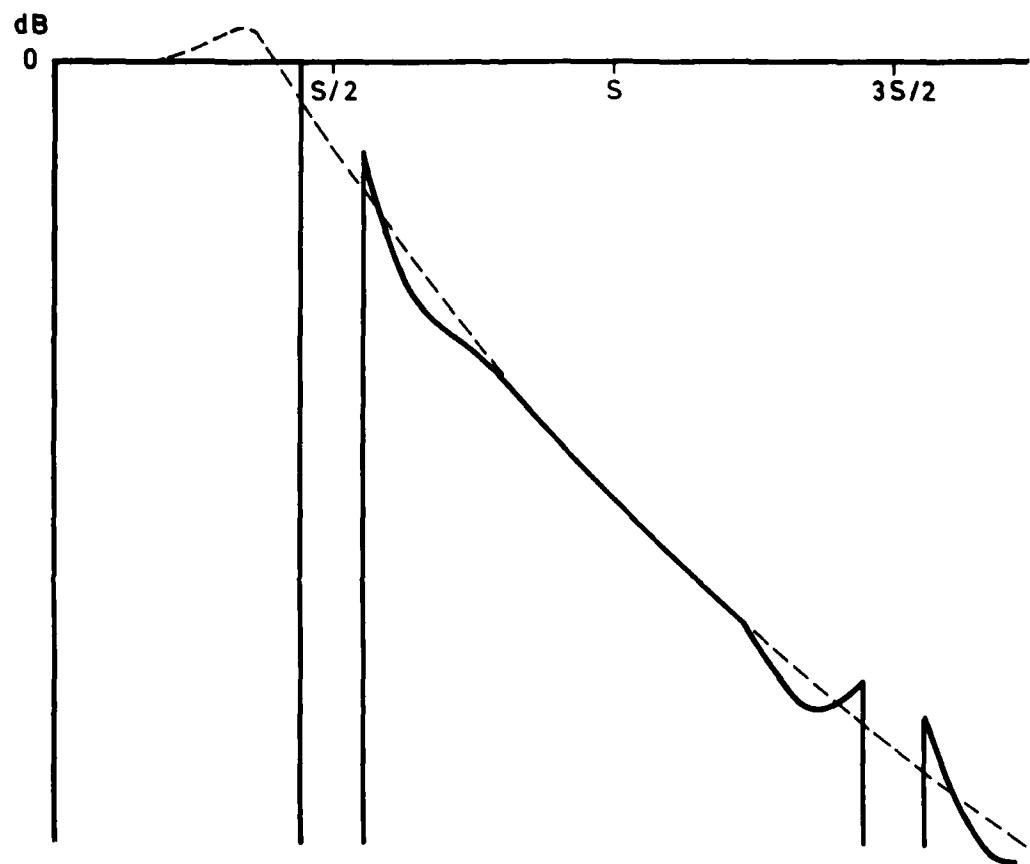


b Effective compensation factor - band limited compensation



c Comparison of raw and compensated response characteristics
-full band compensation

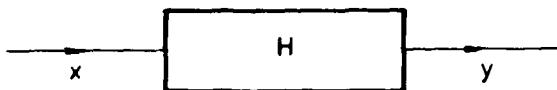
Fig 17d



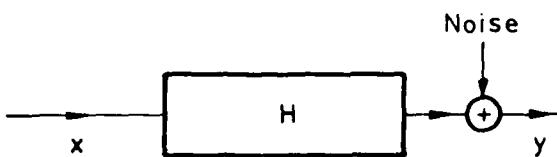
d Comparison of raw and compensated response characteristics
- band limited compensation

Fig 17d Effect of spectral images on effective compensation

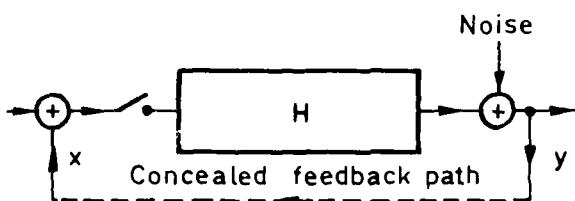
Fig 18a-d



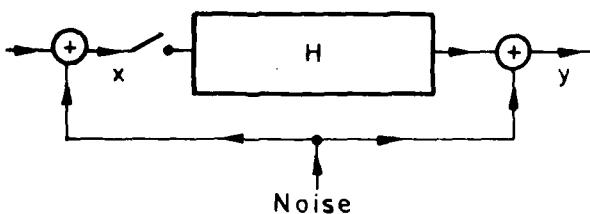
a Single input, single output linear system



b Single input, single output linear system with noise on output



c Feedback can cause 'output' and 'input' to be related, whether or not there is a direct forward transmission path



d Noise which is common to both input and 'output' will cause them to be related, whether or not there is a direct forward transmission path

Fig 18a-d Simple system models

Fig 19

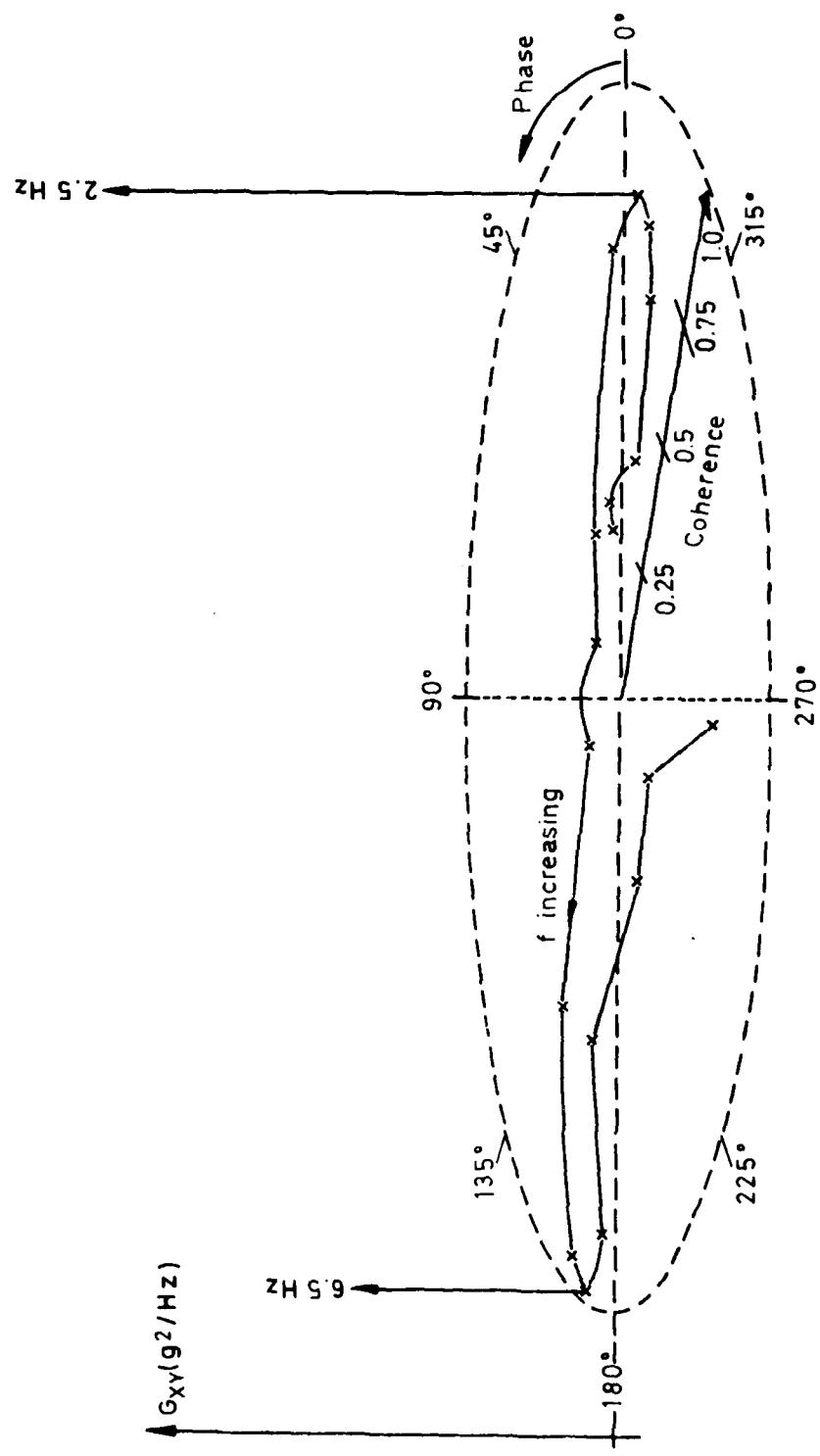


Fig 19 Phase, coherence and cross-spectrum diagram (spectrum computed at 0.5 Hz increments)

REPORT DOCUMENTATION PAGE

Overall security classification of this page

UNLIMITED

As far as possible this page should contain only unclassified information. If it is necessary to enter classified information, the box above must be marked to indicate the classification, e.g. Restricted, Confidential or Secret.

1. DRIC Reference (to be added by DRIC)	2. Originator's Reference RAE TR 81014	3. Agency Reference N/A	4. Report Security Classification/Marking UNLIMITED
5. DRIC Code for Originator 7673000W	6. Originator (Corporate Author) Name and Location Royal Aircraft Establishment, Farnborough, Hants, UK		
5a. Sponsoring Agency's Code N/A	6a. Sponsoring Agency (Contract Authority) Name and Location N/A		
7. Title Digital spectral analysis: A guide based on experience with aircraft vibrations			
7a. (For Translations) Title in Foreign Language			
7b. (For Conference Papers) Title, Place and Date of Conference			
8. Author 1. Surname, Initials Buckingham, S.L.	9a. Author 2 -	9b. Authors 3, 4 -	10. Date February 1981
11. Contract Number N/A	12. Period N/A	13. Project	14. Other Reference Nos. FS 140
15. Distribution statement (a) Controlled by – Unlimited (b) Special limitations (if any) –			
16. Descriptors (Keywords) (Descriptors marked * are selected from TEST) Spectrum analysis*. Fourier transform*. Power spectra*. Digital spectral analysis Fast Fourier transform. Cross spectra. Coherence function. Vibration measurements			
17. Abstract A guide to digital spectral analysis is presented. The emphasis is on a practical engineering understanding of the techniques, based on experience gained in their application to the analysis of flight measurements of aircraft vibration in buffet. Consequently, particular attention is directed to the practical difficulties encountered when the duration of the available data is severely limited, and to the use of the coherence function as a tool in the interpretation of complicated responses. However, the presentation of the fundamental principles, and especially the inherent limitations of the techniques are relevant to any sphere of application.			

1010/1